

A VARIATIONAL TIME DISCRETIZATION FOR COMPRESSIBLE EULER EQUATIONS

FABIO CAVALLETTI, MARC SEDJRO, AND MICHAEL WESTDICKENBERG

ABSTRACT. We introduce a variational time discretization for the multi-dimensional gas dynamics equations, in the spirit of minimizing movements for curves of maximal slope. Each timestep requires the minimization of a functional measuring the acceleration of fluid elements, over the cone of monotone transport maps. We prove convergence to measure-valued solutions for the pressureless gas dynamics and the compressible Euler equations. For one space dimension, we obtain sticky particle solutions for the pressureless case.

1. INTRODUCTION

The compressible Euler equations model the dynamics of compressible fluids like gases. They form a system of hyperbolic conservation laws

$$\left. \begin{aligned} \partial_t \varrho + \nabla \cdot (\varrho \mathbf{u}) &= 0 \\ \partial_t (\varrho \mathbf{u}) + \nabla \cdot (\varrho \mathbf{u} \otimes \mathbf{u}) + \nabla \pi &= 0 \\ \partial_t \varepsilon + \nabla \cdot ((\varepsilon + \pi) \mathbf{u}) &= 0 \end{aligned} \right\} \quad \text{in } [0, \infty) \times \mathbf{R}^d. \quad (1.1)$$

The unknowns $(\varrho, \mathbf{u}, \varepsilon)$ depend on time $t \in [0, \infty)$ and space $x \in \mathbf{R}^d$ and we assume that suitable initial data (to be specified later) is given:

$$(\varrho, \mathbf{u}, \varepsilon)(t = 0, \cdot) =: (\bar{\varrho}, \bar{\mathbf{u}}, \bar{\varepsilon}).$$

We will think of ϱ as a map from $[0, \infty)$ into the space of nonnegative, finite Borel measures, which we denote by $\mathcal{M}_+(\mathbf{R}^d)$. The quantity ϱ is called the density and it represents the distribution of mass in time and space. The first equation in (1.1) (the continuity equation) expresses the local conservation of mass, where

$$\mathbf{u}(t, \cdot) \in \mathcal{L}^2(\mathbf{R}^d, \varrho(t, \cdot)) \quad \text{for all } t \in [0, \infty) \quad (1.2)$$

is the Eulerian velocity field taking values in \mathbf{R}^d . The second equation in (1.1) (the momentum equation) expresses the local conservation of momentum $\mathbf{m} := \varrho \mathbf{u}$. The pressure π will be discussed below. Notice that $\mathbf{m}(t, \cdot)$ is a finite \mathbf{R}^d -valued Borel measure absolutely continuous with respect to $\varrho(t, \cdot)$ for all $t \in [0, \infty)$, because of (1.2). The quantity ε is the total energy of the fluid and $\varepsilon(t, \cdot)$ is again a measure in $\mathcal{M}_+(\mathbf{R}^d)$ for all times $t \in [0, \infty)$. It is reasonable to assume $\varepsilon(t, \cdot)$ to be absolutely continuous with respect to the density $\varrho(t, \cdot)$ (no energy in vacuum). The third (the energy) equation in (1.1) expresses the local conservation of energy.

Date: July 20, 2015.

2000 Mathematics Subject Classification. 35L65, 49J40, 82C40.

Key words and phrases. Compressible Gas Dynamics, Optimal Transport.

Formally, the continuity equation implies that the total mass is preserved:

$$\frac{d}{dt} \int_{\mathbf{R}^d} \varrho(t, dx) = 0 \quad \text{for all } t \in [0, \infty).$$

Similarly, we obtain conservation of the total energy:

$$\frac{d}{dt} \int_{\mathbf{R}^d} \varepsilon(t, dx) = 0 \quad \text{for all } t \in [0, \infty).$$

Therefore, if the fluid has finite mass and total energy initially, then this will also be the case for all positive times. In the following, we will always make this assumption. Without loss of generality, we will also assume that the mass is equal to one, which implies that $\varrho(t, \cdot) \in \mathcal{P}(\mathbf{R}^d)$, the space of Borel probability measures.

To obtain a closed system (1.1) it is necessary to prescribe an equation of state, which relates the pressure π to the density ϱ and the total energy ε . It is provided by thermodynamics. The following three distinct situations are important:

1.1. Pressureless gases. The pressure π vanishes and so the total energy reduces to just the kinetic energy: $\varepsilon = \frac{1}{2} \varrho |\mathbf{u}|^2$. The equations (1.1) take the form

$$\left. \begin{aligned} \partial_t \varrho + \nabla \cdot (\varrho \mathbf{u}) &= 0 \\ \partial_t (\varrho \mathbf{u}) + \nabla \cdot (\varrho \mathbf{u} \otimes \mathbf{u}) &= 0 \end{aligned} \right\} \quad \text{in } [0, \infty) \times \mathbf{R}^d, \quad (1.3)$$

and the energy equation in (1.1) follows formally from the continuity and momentum equations. The system (1.3) has been proposed as a simple model describing the formation of galaxies in the early stage of the universe. Its one-dimensional version is a building block for semiconductor models. Since fluid elements do not interact with each other because there is no pressure, the density $\varrho(t, \cdot)$ may become singular with respect to the d -dimensional Lebesgue measure \mathcal{L}^d . For adhesion (or: sticky particle) dynamics this concentration effect is actually a desired feature; see [65]: If fluid elements meet at the same location, then they stick together to form larger compounds and so $\varrho(t, \cdot)$ can have singular parts (in particular, Dirac measures). Consequently (1.1) must be understood in the sense of distributions. While mass and momentum are conserved, kinetic energy may be destroyed since the collisions are inelastic. In particular, the energy equation in (1.1) will typically be an inequality only. We will call the assumption of adhesion dynamics an entropy condition.

There are now numerous articles studying the pressureless gas dynamics equations (1.3) in one space dimension and establishing global existence of solutions. Frequently, a sequence of approximate solutions is constructed by considering discrete particles, where the initial mass distribution is approximated by a finite sum of Dirac measures. The dynamics of these particles are described by a finite dimensional system of ordinary differential equations between collision times. Whenever multiple particles collide, the new velocity of the bigger particle is determined from the conservation of mass and momentum, and the choice of impact law. The general existence result is obtained by letting the number of discrete particles go to infinity. In order to pass to the limit, several approaches are feasible. We only mention two: One approach relies on the observation that the cumulative distribution function associated to the density ϱ satisfies a certain scalar conservation law (see [13]) so the theory of entropy solutions of scalar conservation laws can be applied. Another approach makes use of the well-known theory of first-order differential inclusions, applied to the cone of monotone transport maps from a reference measure space

to \mathbf{R} ; see [54]. We refer the reader to [8, 9, 12, 34, 38, 42, 44, 53, 55, 56, 63] for more information.

For the multi-dimensional pressureless gas dynamics equations, global existence of solutions to (1.3) is an open problem. The global existence proof in [60] for sticky particle solutions seems to be incomplete, as the authors in [14] show that for a certain choice of initial data, sticky particle solutions cannot exist. This raises the question of the correct solution concept for the equations (1.3).

1.2. Isentropic gases. In this regime, the thermodynamical entropy of the fluid is assumed to be constant in space and time. Consequently, the pressure is a function of the density only. We introduce the internal energy

$$\mathcal{U}[\varrho] := \begin{cases} \int_{\mathbf{R}^d} U(r(x)) dx & \text{if } \varrho = r\mathcal{L}^d, \\ \infty & \text{otherwise,} \end{cases}$$

where $U(r) := \kappa r^\gamma$ for $r \geq 0$. The constant $\gamma > 1$ is called the adiabatic coefficient, and $\kappa > 0$ is another constant. The total energy is the sum of the kinetic energy introduced above and the internal energy. Since we are only interested in solutions of (1.1) with finite total energy, the density $\varrho(t, \cdot)$ must be absolutely continuous with respect to the Lebesgue measure for all $t \in [0, \infty)$. Let $r(t, \cdot)$ be its Radon-Nikodým derivative. Then $\pi(t, \cdot) = P(r(t, \cdot))\mathcal{L}^d$ for all $t \in [0, \infty)$, where

$$P(r) = U'(r)r - U(r) \quad \text{for } r \geq 0.$$

This setup describes polytropic gases. Other choices of U are possible, for example $U(r) = \kappa r \log r$ for isothermal gases (then $P(r) = \kappa r$). We consider

$$\left. \begin{aligned} \partial_t \varrho + \nabla \cdot (\varrho \mathbf{u}) &= 0 \\ \partial_t (\varrho \mathbf{u}) + \nabla \cdot (\varrho \mathbf{u} \otimes \mathbf{u}) + \nabla P(\varrho) &= 0 \end{aligned} \right\} \quad \text{in } [0, \infty) \times \mathbf{R}^d \quad (1.4)$$

(with slight abuse of notation). As in the pressureless case, the energy equation in (1.1) follows formally from the continuity and the momentum equation.

It is well-known that a generic solution to the isentropic Euler equations will not remain smooth, even for regular initial data. Instead the solution will have jump discontinuities along codimension-one submanifolds in space-time, which are called shocks. Then the continuity and the momentum equation must be considered in the sense of distributions, and the energy equation does no longer follow automatically. A physically reasonable relaxation is to assume that no energy can be created by the fluid: The energy equality in (1.1) must be replaced by the inequality

$$\partial_t \left(\frac{1}{2} \varrho |\mathbf{u}|^2 + U(\varrho) \right) + \nabla \cdot \left(\left(\frac{1}{2} \varrho |\mathbf{u}|^2 + U'(\varrho) \varrho \right) \mathbf{u} \right) \leq 0 \quad (1.5)$$

in distributional sense. Physically, strict inequality in (1.5) means that mechanical energy is transformed into heat, a form of energy that is not accounted for by the model. Notice that a differential inequality like (1.5) contains some information on the regularity of solutions: The space-time divergence of a certain nonlinear function of (ϱ, \mathbf{u}) is a nonpositive distribution, and thus a measure. In the one-dimensional case, it is even reasonable to look for weak solutions of (1.4) that satisfy differential inequalities like (1.5) simultaneously for a large class of nonlinear functions of (ϱ, \mathbf{u}) that are called entropy-entropy flux pairs. Such an assumption on the solutions is again an entropy condition. Using the method of compensated

compactness, it is then possible to establish the global existence of weak (entropy) solutions of (1.4). We refer the reader to [18–20, 31–33, 47, 50, 51] for further information. In several space dimensions the only available entropy-entropy flux pair is the total energy-energy flux. Global existence is then an open problem, as is uniqueness of solutions.

1.3. Full Euler equations. We consider a polytropic gas with adiabatic coefficient $\gamma > 1$. Then the pressure is given in terms of $(\varrho, \mathbf{u}, \varepsilon)$ by the formula

$$\pi(t, \cdot) = (\gamma - 1) \left(\varepsilon - \frac{1}{2} \varrho |\mathbf{u}|^2 \right) (t, \cdot) \quad \text{for all } t \in [0, \infty). \quad (1.6)$$

It will be convenient, however, to describe the fluid state using $(\varrho, \mathbf{u}, \sigma)$ instead, where σ denotes the thermodynamical entropy. We assume that $\sigma = \varrho S$, with

$$S(t, \cdot) \in \mathcal{L}^1(\mathbf{R}^d, \varrho(t, \cdot)) \quad \text{for all } t \in [0, \infty)$$

called the specific entropy. For polytropic gases, we again expect that finite energy solutions have a density of the form $\varrho(t, \cdot) = r(t, \cdot) \mathcal{L}^d$. Then we have

$$\pi(t, \cdot) = P(r(t, \cdot), S(t, \cdot)) \mathcal{L}^d \quad \text{for all } t \in [0, \infty), \quad (1.7)$$

where $U(r, S) := \kappa e^S r^\gamma$ for all $r \geq 0$ and $S \in \mathbf{R}$ (with constants $\kappa > 0$ and $\gamma > 1$) and $P(r, S) = U'(r, S)r - U(r, S)$ (the $'$ denoting differentiation with respect to r). Combining (1.6) and (1.7) with (1.1), we formally obtain

$$\partial_t \sigma + \nabla \cdot (\sigma \mathbf{u}) = 0 \quad \text{in } [0, \infty) \times \mathbf{R}^d. \quad (1.8)$$

Equivalently, the specific entropy S is constant along characteristics:

$$\partial_t S + \mathbf{u} \cdot \nabla S = 0 \quad \text{in } [0, \infty) \times \mathbf{R}^d. \quad (1.9)$$

It will be more convenient to work with (1.8) (or even (1.9)) instead of the energy equation in (1.1). Both approaches are formally equivalent. But since solutions to the compressible Euler equations may become discontinuous in finite time, the physically reasonable relaxation is that the specific entropy should only increase when the characteristic crosses a shock forward in time. This restriction is known as the second law of thermodynamics. It follows that

$$\inf_{x \in \mathbf{R}^d} S(t, x) \geq \inf_{x \in \mathbf{R}^d} \bar{S}(x) \quad \text{for all } t \in [0, \infty), \quad (1.10)$$

where \bar{S} is the initial specific entropy. An Eulerian argument in support of (1.10), based on entropy inequalities, was given in [61]. We will assume that

$$\inf_{x \in \mathbf{R}^d} \bar{S}(x) \geq \alpha$$

for some $\alpha \in \mathbf{R}$. Since the shift of S by α can be absorbed into the constant $\kappa > 0$, we may assume without loss of generality that $\alpha = 0$, thus S is nonnegative. As for the isentropic Euler equations, global existence and uniqueness of weak solutions to the full system (1.1) are open problems, now even in one space dimension.

Definition 1.1 (Internal Energy). Let $U(r, S) := \kappa e^S r^\gamma$ for all $r \geq 0$ and $S \in \mathbf{R}$, where $\kappa > 0$ and $\gamma > 1$ are constants. Then we define the internal energy

$$\mathcal{U}[\varrho, \sigma] := \begin{cases} \int_{\mathbf{R}^d} U(r(x), S(x)) \, dx & \text{if } \varrho = r \mathcal{L}^d \text{ and } \sigma = \varrho S, \\ \infty & \text{otherwise,} \end{cases}$$

for all pairs of measures $(\varrho, \sigma) \in \mathcal{P}(\mathbf{R}^d) \times \mathcal{M}_+(\mathbf{R}^d)$.

Weak solutions of hyperbolic conservation laws can be non-unique. Therefore it becomes necessary to augment the conservation laws with an additional assumption, usually called an entropy condition, to select among all possible weak solutions the relevant one. We already mentioned the sticky particle condition for the pressureless case, and the family of differential inequalities for all entropy-entropy flux pairs for the isentropic Euler equations. In either case, the entropy condition is essential for the global existence proofs that are available. But while sticky solutions in one space dimension are unique, the entropy inequalities for the isentropic Euler equations are not sufficient to select a unique solution. In fact, in [23] the authors prove the existence of infinitely many weak solutions of the two-dimensional isentropic Euler equations, starting from the same Lipschitz continuous initial data and satisfying the energy inequality (1.5). The construction uses the technique of convex integration, building on earlier work [25] for the incompressible Euler equations; see also [21, 26]. Notice that for any solution satisfying (1.5), the total energy

$$\mathcal{E}[\varrho, \mathbf{u}](t) := \int_{\mathbf{R}^d} \left(\frac{1}{2} \varrho |\mathbf{u}|^2 + U(\varrho) \right) (t, x) dx \quad \text{for } t \in [0, \infty)$$

is a nonincreasing function in time. Therefore the following entropy condition, called the entropy rate admissibility criterion (see [27]), seems very natural: among all weak solutions of the isentropic Euler equations, pick the one for which $\mathcal{E}[\varrho, \mathbf{u}]$ decreases as fast as possible. It was shown in [22], however, that this entropy condition does not select the solution that seems most physical. More precisely, for the two-dimensional isentropic Euler equations, initial data was given that is constant in x_1 -direction and for which the *expected* solution is self-similar in the x_2 -variable. But then another weak solution was constructed dissipating total energy faster than the self-similar one. This solution is not self-similar, but truly two dimensional.

In this paper, we will consider a different variational principle for the compressible gas dynamics equations that is motivated by minimizing movements for curves of maximal slope on metric spaces; see [4, 30, 45]. Instead of minimizing total energy, we try to decrease the internal energy as fast as possible, while changing the velocity as little as possible. More precisely, we try to find the right balance between dissipating the internal energy and minimizing the acceleration. A similar approach has been used in the context of polyconvex elasticity in [28, 29]. For any given initial data, we define a variational time discretization for (1.1) and prove the convergence of a sequence of approximate solutions to a measure-valued solution of (1.1).

Notice that the concept of measure-valued solutions is rather weak. On the other hand, in view of the non-uniqueness results by De Lellis and Székelyhidi one may wonder whether a distinguished weak solution of (1.1) can be identified at all and what sets it apart from the other solutions. In addition, there is numerical evidence that approximate solutions do not converge as the discretization parameter tends to zero since new features emerge at each length scale. It has therefore been suggested by some researchers that the solution concept for (1.1) must be reconsidered, for example in favor of measure-valued or statistical solutions; see [36, 48, 49].

Here is a short description of our discretization scheme: Recall that in continuum mechanics, a configuration is a function that assigns to each point of the body manifold (the reference configuration) its position in physical space \mathbf{R}^d , at any given time. Typically these maps are required to be injective because matter must not interpenetrate. As a consequence, the space of configurations cannot be a vector space (subtracting a configuration from itself, we obtain the zero map, which is not injective). For our time discretization we adopt a similar point of view: The time interval is partitioned into subintervals of length $\tau > 0$. For each timestep we consider the current fluid state (given by a measure in $\mathcal{P}(\mathbf{R}^{2d})$ that represents the mass distribution of fluid elements and their velocities) as the reference configuration and determine the transport map that moves fluid elements to their new positions. The transport maps are obtained from a suitable minimization problem (see Section 3) and are required to be *monotone*. This ensures that matter does not interpenetrate. Notice, however, that we do not require the maps to be injective (strictly monotone). For the pressureless case, the concentration of mass may be a desired feature of the flow (sticky particles); for the other cases, the invertibility of the transport maps will follow from the presence of pressure. There is no reason why the *global* configuration of the fluid should be monotone (except in one space dimension). But for each step of the time discretization the transport is a perturbation of the identity map, which is monotone. In continuum mechanics, velocities are elements of the tangent space to the manifold of configurations. Therefore, if we consider the map $t \mapsto \varrho(t, \cdot)$ as a curve on the manifold of probability measures, then the corresponding velocity \mathbf{u} should represent a curve in the tangent bundle. In the time discretization we update the velocity as follows: we first move the current velocity using the transport map we just computed, then we project onto a suitably defined tangent cone to the cone of monotone maps at the new configuration. This projection will turn out to be trivial in the cases with pressure; in the pressureless case, the projection of velocity will be related to the sticky particle condition. This two-step update for the velocity is similar to the construction of the parallel transport of tangent vector fields along the space of probability measures, as developed in [3, 41].

Theorem 1.2 (Global Existence). *Let initial data*

$$\bar{\varrho} \in \mathcal{P}_2(\mathbf{R}^d), \quad \bar{\mathbf{u}} \in \mathcal{L}^2(\mathbf{R}^d, \bar{\varrho}), \quad \bar{\sigma} \in \mathcal{M}_+(\mathbf{R}^d)$$

be given such that $\bar{\sigma} = \bar{\varrho} \bar{S}$ with $\bar{S} \in \mathcal{L}^\infty(\mathbf{R}^d, \bar{\varrho})$ nonnegative, and

$$\mathcal{U}[\bar{\varrho}, \bar{\sigma}] < \infty, \quad \int_{\mathbf{R}^d} \bar{\mathbf{u}}(x) \bar{\varrho}(dx) = 0.$$

For any $T > 0$ there exist (see below for the distances involved)

$$\begin{aligned} \varrho &\in \text{Lip}([0, T], \mathcal{P}_2(\mathbf{R}^d)), \quad \sigma \in \text{Lip}([0, T], \mathcal{M}_+(\mathbf{R}^d)), \\ \mathbf{m} &\in \text{Lip}([0, T], \mathcal{M}_K(\mathbf{R}^d)) \end{aligned} \tag{1.11}$$

with the following properties:

- (1) *The initial data is attained:*

$$\varrho(0, \cdot) = \bar{\varrho}, \quad \sigma(0, \cdot) = \bar{\sigma}, \quad \mathbf{m}(0, \cdot) = \bar{\varrho} \bar{\mathbf{u}}.$$

- (2) *We have $\mathbf{m} =: \varrho \mathbf{u}$ with $\mathbf{u}(t, \cdot) \in \mathcal{L}^2(\mathbf{R}^d, \varrho(t, \cdot))$ for all $t \in [0, T]$.*

- (3) *There exist suitable Young measures (see Section 6) that generate*

- the momentum flux $\langle \varrho \mathbf{u} \otimes \mathbf{u} \rangle \in \mathcal{L}_w^\infty([0, T], \mathcal{M}(\mathbf{R}^d; \mathcal{S}_+^d))$,
- the pressure tensor $\langle \boldsymbol{\pi} \rangle \in \mathcal{L}_w^\infty([0, T], \mathcal{M}(\mathbf{R}^d; \mathcal{S}_+^d))$,
- the entropy flux $\langle \sigma \mathbf{u} \rangle \in \mathcal{L}_w^\infty([0, T], \mathcal{M}(\mathbf{R}^d; \mathbf{R}^d))$,

such that

$$\left. \begin{aligned} \partial_t \varrho + \nabla \cdot (\varrho \mathbf{u}) &= 0 \\ \partial_t (\varrho \mathbf{u}) + \nabla \cdot \langle \varrho \mathbf{u} \otimes \mathbf{u} \rangle + \nabla \cdot \langle \boldsymbol{\pi} \rangle &= 0 \\ \partial_t \sigma + \nabla \cdot \langle \sigma \mathbf{u} \rangle &= 0 \end{aligned} \right\} \quad \text{in } \mathcal{D}'([0, \infty) \times \mathbf{R}^d). \quad (1.12)$$

We denote by $\mathcal{P}_2(\mathbf{R}^d)$ the space of Borel probability measures with finite second moment, endowed with the 2-Wasserstein distance; see Definition 2.1 below. Since σ satisfies a transport equation, the total entropy is preserved, so we can define its Lipschitz continuity again with respect to a 2-Wasserstein distance. We denote by $\mathcal{M}_K(\mathbf{R}^d)$ the space of \mathbf{R}^d -valued Borel measures \mathbf{m} with zero mean and finite first moment, equipped with the Monge-Kantorovich norm (see [24])

$$\|\mathbf{m}\|_{\mathcal{M}_K(\mathbf{R}^d)} := \sup \left\{ \int_{\mathbf{R}^d} \zeta(x) \cdot \mathbf{m}(dx) : \|\zeta\|_{\text{Lip}(\mathbf{R}^d)} \leq 1 \right\}, \quad (1.13)$$

where the Lipschitz seminorm of a function $\zeta : \mathbf{R}^d \rightarrow \mathbf{R}^d$ is defined as

$$\|\zeta\|_{\text{Lip}(\mathbf{R}^d)} := \sup_{x \neq y} \frac{|\zeta(x) - \zeta(y)|}{|x - y|}.$$

Since $|\zeta(x)| \leq |\zeta(0)| + |x|$ for all $x \in \mathbf{R}^d$ when $\|\zeta\|_{\text{Lip}(\mathbf{R}^d)} \leq 1$, the integral in (1.13) is finite if $\mathbf{m} = \varrho \mathbf{u}$ with $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ and $\mathbf{u} \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$, by Cauchy-Schwarz.

We assume in Theorem 1.2 that the total momentum vanishes initially (which implies that the total momentum vanishes for all $t \in [0, T]$). This is not a restriction since the hyperbolic conservation law (1.1) is invariant under transformations to a moving reference frame (in the absence of boundaries). For the proof of our result, we use Young measures on the space of (generalized) curves on \mathbf{R}^d . In particular, we will be able to capture the (superposition of) trajectories along which a fluid element travels that was located at a generic position $\bar{x} \in \mathbf{R}^d$ at time $t = 0$. We will show that mass and entropy travel along the same trajectories (see (6.21)/(6.22)), so that the entropy flux $\langle \sigma \mathbf{u} \rangle$ and $\varrho \mathbf{u}$ are related: Both can be computed from the same Young measure. If the Young measure becomes trivial in the sense that each fluid element travels along a single trajectory that does not intersect the path of any other fluid element, then $\langle \sigma \mathbf{u} \rangle = \sigma \mathbf{u}$. We also expect the pressure tensor $\langle \boldsymbol{\pi} \rangle$ to be diagonal and determined by a scalar pressure that is a function of ϱ and σ . Our discretization generates an additional stress tensor field that serves as a Lagrange multiplier for the monotonicity constraint we impose on the transport maps in each timestep. This component does not appear in the momentum equation.

Our variational time discretization decreases the total energy, while preserving the entropy. This may seem backwards from the physical point of view. We would like to point out, however, that in turbulence it is standard to assume that solutions of the incompressible Navier-Stokes equations converge (in the high Reynolds number limit) to velocity fields that dissipate kinetic energy, even though they formally solve the incompressible Euler equations. Therefore the incompressible Euler equations seem to only give an incomplete description of the actual physical phenomena. It is natural to expect that similar effects occur in the compressible models.

2. NOTATION

In the following, we will always assume that \mathbf{R}^D is equipped with the Euclidean inner product, for which we write $x \cdot y =: \langle x, y \rangle$ with $x, y \in \mathbf{R}^D$. For matrices we will use the inner product $\text{tr}(A^T B) =: \langle\langle A, B \rangle\rangle$ for all $A, B \in \mathbf{R}^{D \times D}$. The norms on these spaces will always be the ones induced by the inner products.

We denote by $\mathcal{C}_b(\mathbf{R}^D)$ the space of bounded continuous functions on \mathbf{R}^D and by $\mathcal{P}(\mathbf{R}^D)$ the space of Borel probability measures. Weak convergence of sequences of probability measures is defined by testing against functions in $\mathcal{C}_b(\mathbf{R}^D)$. For any $1 \leq p < \infty$ we denote by $\mathcal{P}_p(\mathbf{R}^D)$ the space of Borel probability measures with finite p th moment, so that $\int_{\mathbf{R}^D} |x|^p \varrho(dx) < \infty$ for every $\varrho \in \mathcal{P}_p(\mathbf{R}^D)$.

Definition 2.1 (p -Wasserstein Distance). For any $\varrho^1, \varrho^2 \in \mathcal{P}(\mathbf{R}^D)$ let

$$\text{ADM}(\varrho^1, \varrho^2) := \left\{ \gamma \in \mathcal{P}(\mathbf{R}^{2D}) : \mathbb{P}^k \# \gamma = \varrho^k \text{ with } k = 1..2 \right\}$$

be the space of admissible transport plans connecting ϱ^1 and ϱ^2 , where

$$\mathbb{P}^k(x^1, x^2) := x^k \quad \text{for all } (x^1, x^2) \in \mathbf{R}^{2D} = (\mathbf{R}^D)^2$$

and $k = 1..2$, and $\#$ denotes the push-forward of measures. For any $1 \leq p < \infty$ the p -Wasserstein distance $W_p(\varrho^1, \varrho^2)$ between ϱ^1, ϱ^2 is defined by

$$W_p(\varrho^1, \varrho^2)^p := \inf_{\gamma \in \text{ADM}(\varrho^1, \varrho^2)} \int_{\mathbf{R}^{2D}} |x^1 - x^2|^p \gamma(dx^1, dx^2). \quad (2.1)$$

It is well-known that the inf in (2.1) is actually attained, so the set $\text{OPT}(\varrho^1, \varrho^2)$ of transport plans γ that minimize (2.1) (called optimal transport plans) is nonempty. For $p = 2$ the support of each $\gamma \in \text{OPT}(\varrho^1, \varrho^2)$ is contained in the subdifferential of a lower semicontinuous, convex map (therefore it is cyclically monotone). If ϱ^1 is absolutely continuous with respect to the Lebesgue measure \mathcal{L}^D , then each optimal transport plan is induced by a map (its support lies on the graph of a function):

$$\gamma = (\text{id}, \mathcal{T}) \# \varrho^1 \quad \text{for suitable } \mathcal{T} \in \mathcal{L}^2(\mathbf{R}^D, \varrho^1).$$

We refer the reader to [4] for further details.

For any $n \in \mathbf{N}$ and $k = 1 \dots n$, we define projections

$$\mathbb{P}^k(x^1 \dots x^n) := x^k \quad \text{for all } (x^1 \dots x^n) \in \mathbf{R}^{nd} = (\mathbf{R}^d)^n.$$

We will also use projections \mathbb{x} and \mathbb{y}^k defined by

$$\mathbb{x}(x, y^1 \dots y^n) := x, \quad \mathbb{y}^k(x, y^1 \dots y^n) := y^k$$

for all $(x, y^1 \dots y^n) \in \mathbf{R}^{(n+1)d} = (\mathbf{R}^d)^{n+1}$ and $k = 1 \dots n$, with $n \in \mathbf{N}$. Sometimes it will be convenient to write \mathbb{z}^k or \mathbb{w}^k in place of \mathbb{y}^k (same definition), depending on whether the symbols represent positions or velocities, which will be clear from the context. For $n = 1$ we will usually write $\mathbb{y} := \mathbb{y}^1$ etc.

Definition 2.2 (Distance). Let $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ be given and

$$\mathcal{P}_\varrho(\mathbf{R}^{2d}) := \left\{ \gamma \in \mathcal{P}_2(\mathbf{R}^{2d}) : \mathbb{x} \# \gamma = \varrho \right\}.$$

We introduce a distance as follows: for any $\gamma^1, \gamma^2 \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ we define

$$W_\varrho(\gamma^1, \gamma^2)^2 := \int_{\mathbf{R}^d} W(\gamma_x^1, \gamma_x^2)^2 \varrho(dx),$$

where $\gamma^k(dx, dy) =: \gamma_x^k(dy) \varrho(dx)$ with $k = 1..2$ denotes the disintegration of γ^k , and where W is the Wasserstein distance on $\mathcal{P}_2(\mathbf{R}^d)$; see [4, 41].

Definition 2.3 (Transport Plans). Let $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ be given.

(i.) *Admissible Plans.* For any $\gamma^1, \gamma^2 \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ we define

$$\text{ADM}_\varrho(\gamma^1, \gamma^2) := \left\{ \alpha \in \mathcal{P}_2(\mathbf{R}^{3d}) : (\mathbb{x}, \mathbb{y}^k) \# \alpha = \gamma^k \text{ with } k = 1..2 \right\}.$$

(ii.) *Optimal Plans.* For any $\gamma^1, \gamma^2 \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ we define

$$\begin{aligned} \text{OPT}_\varrho(\gamma^1, \gamma^2) &:= \left\{ \alpha \in \text{ADM}_\varrho(\gamma^1, \gamma^2) : \right. \\ &\quad \left. W_\varrho(\gamma^1, \gamma^2)^2 = \int_{\mathbf{R}^{3d}} |y^1 - y^2|^2 \alpha(dx, dy^1, dy^2) \right\}. \end{aligned}$$

Theorem 2.4. Let $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ be given.

(i.) The function W_ϱ is a distance on $\mathcal{P}_\varrho(\mathbf{R}^{2d})$ and lower semicontinuous with respect to weak convergence in $\mathcal{P}_2(\mathbf{R}^{2d})$. We have

$$W_\varrho(\gamma^1, \gamma^2)^2 = \min_{\alpha \in \text{ADM}_\varrho(\gamma^1, \gamma^2)} \int_{\mathbf{R}^{3d}} |y^1 - y^2|^2 \alpha(dx, dy^1, dy^2)$$

for all $\gamma^1, \gamma^2 \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$, and thus $\text{OPT}_\varrho(\gamma^1, \gamma^2)$ is nonempty.

(ii.) The set $(\mathcal{P}_\varrho(\mathbf{R}^{2d}), W_\varrho)$ is a complete metric space.

Proof. We refer the reader to Section 4.1 in [41]. \square

Definition 2.5 (Barycentric Projection). For any $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ and $\gamma \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ the barycentric projection $\mathbb{B}(\gamma) \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$ is defined as

$$\mathbb{B}(\gamma)(x) := \int_{\mathbf{R}^d} y \gamma_x(dy) \quad \text{for } \varrho\text{-a.e. } x \in \mathbf{R}^d,$$

where $\gamma(dx, dy) =: \gamma_x(dy) \varrho(dx)$ is the disintegration of γ .

An important subset of $\mathcal{P}_\varrho(\mathbf{R}^{2d})$ consists of those measures γ that are induced by maps: there exists a $\mathcal{Y} \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$ taking values in \mathbf{R}^d such that

$$\gamma(dx, dy) = \delta_{\mathcal{Y}(x)}(dy) \varrho(dx).$$

In this case, the distance W_ϱ reduces to the $\mathcal{L}^2(\mathbf{R}^d, \varrho)$ -distance of the corresponding maps. If $\gamma^1, \gamma^2 \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ and $\gamma^1 = (\text{id}, \mathcal{Y}^1) \# \varrho$ with $\mathcal{Y}^1 \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$, then

$$W_\varrho(\gamma^1, \gamma^2)^2 = \int_{\mathbf{R}^{2d}} |\mathcal{Y}^1(x) - y^2|^2 \gamma^2(dx, dy^2); \quad (2.2)$$

see Lemma 5.3.2 in [4]. Hence, if $W_\varrho(\gamma^n, \gamma) \rightarrow 0$ as $n \rightarrow \infty$, with $\gamma^n, \gamma \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ and $\gamma^n = (\text{id}, \mathcal{Y}^n) \# \varrho$ for suitable $\mathcal{Y}^n \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$, then $\mathcal{Y}^n \rightarrow \mathcal{Y}^\infty$ strongly in $\mathcal{L}^2(\mathbf{R}^d, \varrho)$ and $\gamma = (\text{id}, \mathcal{Y}^\infty) \# \varrho$. Indeed, the assumption implies that $\{\mathcal{Y}^n\}_n$ is a Cauchy sequence in $\mathcal{L}^2(\mathbf{R}^d, \varrho)$ and hence converges to a limit \mathcal{Y}^∞ , by completeness. Since $(\text{id}, \mathcal{Y}^n, \mathcal{Y}^\infty) \# \varrho \in \text{ADM}_\varrho(\gamma^n, \gamma^\infty)$ with $\gamma^\infty := (\text{id}, \mathcal{Y}^\infty) \# \varrho$, we have

$$W_\varrho(\gamma^n, \gamma^\infty) \leq \|\mathcal{Y}^n - \mathcal{Y}^\infty\|_{\mathcal{L}^2(\mathbf{R}^d, \varrho)} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Then we use that $W_\varrho(\gamma, \gamma^\infty) \leq W_\varrho(\gamma^n, \gamma) + W_\varrho(\gamma^n, \gamma^\infty)$. We have the estimate

$$\|\mathbb{B}(\gamma^1) - \mathbb{B}(\gamma^2)\|_{\mathcal{L}^2(\mathbf{R}^d, \varrho)} \leq W_\varrho(\gamma^1, \gamma^2),$$

as follows easily from Theorem 2.4 (i.) and Jensen inequality.

3. ENERGY MINIMIZATION: FIRST PROPERTIES

In preparation of our variational time discretization for (1.1), we first consider the metric projection onto the cone of monotone transport plans.

3.1. Monotone Transport Plans. To every subset $\Gamma \subset \mathbf{R}^d \times \mathbf{R}^d$ we can associate a set-valued map $u_\Gamma: \mathbf{R}^d \longrightarrow P(\mathbf{R}^d)$ (where $P(\mathbf{R}^d)$ is the power set of \mathbf{R}^d) by

$$u_\Gamma(x) := \left\{ y \in \mathbf{R}^d : (x, y) \in \Gamma \right\} \quad \text{for all } x \in \mathbf{R}^d.$$

For any set-valued map $u: \mathbf{R}^d \longrightarrow P(\mathbf{R}^d)$, we denote by

$$\begin{aligned} \text{dom}(u) &:= \left\{ x \in \mathbf{R}^d : u(x) \neq \emptyset \right\}, \\ \text{graph}(u) &:= \left\{ (x, y) \in \mathbf{R}^d \times \mathbf{R}^d : y \in u(x) \right\} \end{aligned}$$

its domain and graph. A subset $\Gamma \subset \mathbf{R}^d \times \mathbf{R}^d$ is called monotone if

$$\langle x_1 - x_2, y_1 - y_2 \rangle \geq 0 \quad \text{for any pair of } (x_i, y_i) \in \Gamma.$$

Such a set is called maximal monotone if for any monotone set $\Gamma' \subset \mathbf{R}^d \times \mathbf{R}^d$ with $\Gamma \subset \Gamma'$ we have that $\Gamma = \Gamma'$. Equivalently, if it is not possible to enlarge Γ without destroying the monotonicity. We will call any set-valued map u as above (maximal) monotone if the set $\text{graph}(u)$ is (maximal) monotone.

By Zorn's lemma, any monotone set (equivalently, any monotone set-valued map) can be extended to a maximal monotone set (map). Typically, this extension is not unique. A maximal monotone extension can be obtained constructively as follows: Let $\Gamma \subset \mathbf{R}^d \times \mathbf{R}^d$ be monotone. Then (for all $(x, y), (x^*, y^*) \in \mathbf{R}^d \times \mathbf{R}^d$)

- (1) define the Fitzpatrick function

$$F_\Gamma(x, y) := \sup \left\{ \langle y', x \rangle + \langle y, x' \rangle - \langle y', x' \rangle : (x', y') \in \Gamma \right\};$$

- (2) compute its Fenchel conjugate

$$F_\Gamma^*(y^*, x^*) := \sup \left\{ \langle y^*, x \rangle + \langle y, x^* \rangle - F_\Gamma(x, y) : (x, y) \in \mathbf{R}^d \times \mathbf{R}^d \right\};$$

- (3) compute the proximal average

$$\begin{aligned} N_\Gamma(x, y) &:= \inf \left\{ \frac{1}{2} F_\Gamma(x_1, y_1) + \frac{1}{2} F_\Gamma^*(y_2, x_2) + \frac{1}{8} \|x_1 - x_2\|^2 + \frac{1}{8} \|y_1 - y_2\|^2 : \right. \\ &\quad \left. (x, y) = \frac{1}{2}(x_1, y_1) + \frac{1}{2}(x_2, y_2) \right\}. \end{aligned}$$

The function N_Γ is lower semicontinuous, convex, and proper, and the set

$$\bar{\Gamma} := \left\{ (x, y) : N_\Gamma(x, y) = \langle y, x \rangle \right\} \tag{3.1}$$

is a maximal monotone extension of Γ . We refer the reader to [5, 40] for details.

Remark 3.1. For any maximal monotone set-valued function $u: \mathbf{R}^d \longrightarrow P(\mathbf{R}^d)$ the image $u(x)$ of any $x \in \mathbf{R}^d$ is closed and convex (possibly empty); see Proposition 1.2 of [1]. Therefore the dimension $\dim u(x)$ is well-defined. The singular sets

$$\Sigma^k(u) := \left\{ x \in \mathbf{R}^d : \dim u(x) \geq k \right\}, \quad \text{with } k = 1 \dots d,$$

are countably \mathcal{H}^{d-k} -rectifiable; see Theorem 2.2 of [1] for details. Here \mathcal{H}^n denotes the n -dimensional Hausdorff measure. In particular, the set of points $x \in \text{dom}(u)$

for which $u(x)$ contains more than one point (that is, the set $\Sigma^1(u)$) is negligible with respect to the Lebesgue measure \mathcal{L}^d . Outside $\Sigma^1(u)$ the function u is continuous. This observation will allow us to think of a maximal monotone map u as a Lebesgue measurable, *single-valued* function (just redefine u on the null set $\Sigma^1(u)$).

Definition 3.2 (Monotone Transport Plans). For any $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$, we define

$$C_\varrho := \left\{ \gamma \in \mathcal{P}_\varrho(\mathbf{R}^{2d}) : \text{spt } \gamma \text{ is a monotone subset of } \mathbf{R}^d \times \mathbf{R}^d \right\}. \quad (3.2)$$

Our definition of monotonicity for measures in $\mathcal{P}_\varrho(\mathbf{R}^{2d})$ is motivated by the optimal transport plans of Definition 2.1: an *optimal* transport plan γ is characterized by the property that $\text{spt } \gamma$ must be a *cyclically monotone* set; see Section 6.2.3 in [4]. Then there exists a lower semicontinuous, convex, proper function φ with

$$\varphi(x) + \varphi^*(y) = \langle y, x \rangle \quad \text{for } \gamma\text{-a.e. } (x, y) \in \mathbf{R}^d \times \mathbf{R}^d.$$

Here φ^* denotes the Fenchel conjugate to φ . In our setting, the cyclical monotonicity is replaced by monotonicity, and $N_\Gamma(x, y)$ of (3.1) plays the role of $\varphi(x) + \varphi^*(y)$. In the terminology of [40], the function N_Γ is called a self-dual Lagrangian. The cone C_ϱ contains the set of optimal transport plans defined above.

Since we do not make any assumptions on ϱ , its support may be a proper subset of \mathbf{R}^d and have “holes”. Fortunately, the monotonicity constraint enables us to work with objects that are defined on a fixed convex open subset of \mathbf{R}^d :

Definition 3.3 (Associated Maps). Let $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ be given. For every $\gamma \in C_\varrho$ we call u a maximal monotone map associated to γ if u is the maximal monotone set-valued map induced by a maximal monotone extension of $\Gamma := \text{spt } \gamma$.

Lemma 3.4. For any given $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ and $\gamma \in C_\varrho$, the domain of every maximal monotone map u associated to γ contains the convex open set $\Omega := \text{int } \overline{\text{conv}} \text{spt } \varrho$, where int denotes the interior and $\overline{\text{conv}}$ is the closed convex hull.

Proof. Let $\gamma \in C_\varrho$ be given and consider any maximal monotone map u associated to γ . Then $\text{graph}(u)$ is a maximal monotone extension of $\Gamma := \text{spt } \gamma$, which implies that the projection $X := \text{p}^1(\Gamma)$ of Γ onto \mathbf{R}^d is contained in $\text{dom}(u)$. Since

$$\text{int } \overline{\text{conv}} \text{dom}(u) \subset \text{dom}(u) \subset \overline{\text{conv}} \text{dom}(u)$$

(this is true for every maximal monotone set-valued function; see Corollary 1.3 of [1]) we conclude that the convex open set $\text{int } \overline{\text{conv}}(X) \subset \text{dom}(u)$. It therefore remains to show that $\text{int } \overline{\text{conv}}(X) = \Omega$. Note that Ω is independent of γ and u .

To prove the claim, choose any $x \in X$ and $r > 0$. Then we can estimate

$$\varrho(B_r(x)) = \gamma(B_r(x) \times \mathbf{R}^d) \geq \gamma(B_r(x) \times B_r(y)) > 0,$$

for suitable $y \in \mathbf{R}^d$ with $(x, y) \in \Gamma = \text{spt } \gamma$. Since $x \in X$ and $r > 0$ were arbitrary, we get that $X \subset \text{spt } \varrho$, which implies that $\text{int } \overline{\text{conv}}(X) \subset \Omega$.

Conversely, for every $x \in \Omega$ there exists a ball $B_r(x) \subset \overline{\text{conv}} \text{spt } \varrho$ for some $r > 0$. Pick an open d -cube Q centered at x as large as possible with $Q \subset B_{r/2}(x)$. Then the closure \overline{Q} is the convex hull of its corners $x_i \in \partial B_{r/2}(x)$, which satisfy

$$x_i \in \overline{\text{conv}} \text{spt } \varrho \quad \text{for } i = 1 \dots 2^d.$$

Let $\ell > 0$ denote the side length of Q and $0 < \varepsilon < \ell/8$. Then there exist

$$y_i \in B_\varepsilon(x_i) \cap \text{conv spt } \varrho \quad \text{for } i = 1 \dots 2^d.$$

Each y_i can be written as a convex combination

$$y_i = \sum_{k=1}^{N_i} \lambda_{i,k} z_{i,k} \quad \text{with } \lambda_{i,k} \in [0, 1] \text{ and } \sum_{k=1}^{N_i} \lambda_{i,k} = 1,$$

for suitable $z_{i,k} \in \text{spt } \varrho$ and $N_i \in \mathbf{N}$. We now claim that for any $z \in \text{spt } \varrho$ and $\varepsilon > 0$ there exists $\bar{z} \in B_\varepsilon(z) \cap X$. Assume for the moment that the claim is true. Then for each $z_{i,k}$ we can find $\bar{z}_{i,k} \in B_\varepsilon(z_{i,k}) \cap X$. We define convex combinations

$$\bar{y}_i := \sum_{k=1}^{N_i} \lambda_{i,k} \bar{z}_{i,k} \quad \text{for all } i = 1 \dots 2^d,$$

which satisfy $\|y_i - \bar{y}_i\| \leq \varepsilon$ and thus $\bar{y}_i \in B_{2\varepsilon}(x_i)$ for all i . Consequently, the convex hull of these \bar{y}_i contains a d -cube centered at x with side length $\ell/2$, which in turn contains a ball $B_\delta(x)$ for $\delta > 0$ small enough. By construction, this ball is a subset of the convex hull of the $\bar{z}_{i,k} \in X$ from above, so that $x \in \text{int } \overline{\text{conv}}(X)$. This proves the lemma. To establish the claim, assume that on the contrary, there exists $\varepsilon > 0$ with the property that for all $\bar{z} \in B_\varepsilon(z)$ we have $\bar{z} \notin X$. Then

$$\begin{aligned} \varrho(B_\varepsilon(z)) &= \gamma(B_\varepsilon(z) \times \mathbf{R}^d) \\ &\leq \gamma((\mathbf{R}^d \setminus X) \times \mathbf{R}^d) \leq \gamma((\mathbf{R}^d \times \mathbf{R}^d) \setminus \Gamma) = 0. \end{aligned}$$

The second equality follows from the fact that γ (being a finite Borel measure on a locally compact Hausdorff space with countable basis) is inner regular; see [37]. We conclude that $z \notin \text{spt } \varrho$, which is a contradiction. \square

3.2. Minimal Acceleration Cost. Any plan $\mu \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ describes the state of some fluid with density $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ and velocity distribution μ_x for ϱ -a.e. $x \in \mathbf{R}^d$, where $\mu(dx, d\xi) =: \mu_x(d\xi) \varrho(dx)$ denotes the disintegration of μ with respect to ϱ . The special case $\mu(dx, d\xi) = \delta_{u(x)}(d\xi) \varrho(dx)$ represents a monokinetic state where all fluid elements located at the position $x \in \mathbf{R}^d$ have the same velocity $u(x) \in \mathbf{R}^d$ and are therefore indistinguishable. The velocity field $u \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$, by construction. We will occasionally use bold letters to denote elements in \mathbf{R}^{2d} such as

$$\mathbf{x} = (x, \xi), \quad \mathbf{y} = (y, v), \quad \text{and} \quad \mathbf{z} = (z, \zeta),$$

where $x, y, z \in \mathbf{R}^d$ represent positions and $\xi, v, \zeta \in \mathbf{R}^d$ velocities.

In order to measure the “distance” between two state measures $\mu^1, \mu^2 \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$, we will use the minimal acceleration cost introduced in [39]. It is defined as follows: For a given timestep $\tau > 0$, consider a fluid element with initial position/velocity $\mathbf{x} \in \mathbf{R}^{2d}$. Assume that the fluid element transitions into a new state $\mathbf{z} \in \mathbf{R}^{2d}$. The transition is described by a smooth curve $X(\cdot | \mathbf{x}, \mathbf{z}): [0, \tau] \rightarrow \mathbf{R}^{2d}$ such that

$$(X, \dot{X})(0) = (x, \xi) \quad \text{and} \quad (X, \dot{X})(\tau) = (z, \zeta)$$

(with $X := X(\cdot | \mathbf{x}, \mathbf{z})$). Among all such curves there are the ones that minimize the acceleration $\int_0^\tau |\dot{X}(t)|^2 dt$. They are uniquely determined and given by

$$X(t | \mathbf{x}, \mathbf{z}) = x + t\xi + \left(3(z - x) - \tau(\zeta + 2\xi)\right) \frac{t^2}{\tau^2} - \left(2(z - x) - \tau(\zeta + \xi)\right) \frac{t^3}{\tau^3}$$

for all $t \in [0, \tau]$. The minimal acceleration can be computed explicitly, which allows us to define a cost measuring the “distance” between the two end states:

$$a_\tau(\mathbf{x}, \mathbf{z})^2 := 3 \left| \frac{\mathbf{z} - \mathbf{x}}{\tau} - \frac{\boldsymbol{\zeta} + \boldsymbol{\xi}}{2} \right|^2 + \frac{1}{4} |\boldsymbol{\zeta} - \boldsymbol{\xi}|^2 \quad (3.3)$$

for all $\mathbf{x}, \mathbf{z} \in \mathbf{R}^{2d}$. Note that $a_\tau(\mathbf{x}, \mathbf{z}) = 0$ if and only if $\mathbf{z} = \mathbf{x} + \tau\boldsymbol{\xi}$ and $\boldsymbol{\zeta} = \boldsymbol{\xi}$.

Definition 3.5 (Minimal Acceleration Cost). For all measures $\boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in \mathcal{P}_2(\mathbf{R}^{2d})$ let $\text{ADM}(\boldsymbol{\mu}^1, \boldsymbol{\mu}^2)$ denote the set of transport plans $\boldsymbol{\omega} \in \mathcal{P}(\mathbf{R}^{4d})$ with

$$(\mathbb{P}^1, \mathbb{P}^2) \# \boldsymbol{\omega} = \boldsymbol{\mu}^1 \quad \text{and} \quad (\mathbb{P}^3, \mathbb{P}^4) \# \boldsymbol{\omega} = \boldsymbol{\mu}^2.$$

The minimal acceleration cost is the functional A_τ defined by

$$A_\tau(\boldsymbol{\mu}^1, \boldsymbol{\mu}^2)^2 := \inf \left\{ \int_{\mathbf{R}^{4d}} a_\tau(\mathbf{x}^1, \mathbf{x}^2)^2 \boldsymbol{\omega}(d\mathbf{x}^1, d\mathbf{x}^2) : \boldsymbol{\omega} \in \text{ADM}(\boldsymbol{\mu}^1, \boldsymbol{\mu}^2) \right\}. \quad (3.4)$$

Note that A_τ is not a distance: It is not symmetric in its arguments $\boldsymbol{\mu}^1$ and $\boldsymbol{\mu}^2$, which follows from the asymmetry of the cost function (3.3). Moreover, it does not vanish if $\boldsymbol{\mu}^1 = \boldsymbol{\mu}^2$. Instead, we have the following relation:

$$A_\tau(\boldsymbol{\mu}^1, \boldsymbol{\mu}^2) = 0 \quad \Longleftrightarrow \quad \boldsymbol{\mu}^2 = F_\tau \# \boldsymbol{\mu}^1,$$

where $F_\tau : \mathbf{R}^{2d} \rightarrow \mathbf{R}^{2d}$ is the *free transport map* defined by

$$F_\tau(\mathbf{x}) := (\mathbf{x} + \tau\boldsymbol{\xi}, \boldsymbol{\xi}) \quad \text{for all } \mathbf{x} \in \mathbf{R}^{2d}.$$

The minimal acceleration cost measures how much each fluid element deviates from the straight path determined by its initial velocity; see [39] for more details.

The cost function (3.3) can be rewritten in the following form:

$$a_\tau(\mathbf{x}, \mathbf{z})^2 = \frac{3}{4\tau^2} |(\mathbf{x} + \tau\boldsymbol{\xi}) - \mathbf{z}|^2 + \left| \boldsymbol{\zeta} - \left(\boldsymbol{\xi} - \frac{3}{2\tau} ((\mathbf{x} + \tau\boldsymbol{\xi}) - \mathbf{z}) \right) \right|^2 \quad (3.5)$$

for every $\mathbf{x}, \mathbf{z} \in \mathbf{R}^{2d}$. The first term measures how much the final position \mathbf{z} differs from $\mathbf{x} + \tau\boldsymbol{\xi}$, which would be the position of the fluid element after a free transport. The second term measures the difference between $\boldsymbol{\zeta}$ and the velocity

$$V_\tau(\mathbf{x}, \mathbf{z}) := \boldsymbol{\xi} - \frac{3}{2\tau} ((\mathbf{x} + \tau\boldsymbol{\xi}) - \mathbf{z}), \quad (3.6)$$

which is the velocity that minimizes the acceleration among all curves that connect the initial position/velocity $\mathbf{x} \in \mathbf{R}^{2d}$ to the final position $\mathbf{z} \in \mathbf{R}^d$. Consequently, when minimizing the integral in (3.4) over all plans $\boldsymbol{\omega} \in \mathcal{P}_2(\mathbf{R}^{4d})$ with

$$(\mathbb{P}^1, \mathbb{P}^2, \mathbb{P}^3) \# \boldsymbol{\omega} =: \boldsymbol{\beta} \quad \text{for given } \boldsymbol{\beta} \in \mathcal{P}_2(\mathbf{R}^{3d}),$$

then there exists a unique such minimizer, which takes the form $\boldsymbol{\omega} = H_\tau \# \boldsymbol{\beta}$, with the map $H_\tau : \mathbf{R}^{3d} \rightarrow \mathbf{R}^{4d}$ defined for all $\mathbf{x} \in \mathbf{R}^{2d}$ and $\mathbf{z} \in \mathbf{R}^d$ as

$$H_\tau(\mathbf{x}, \mathbf{z}) := (\mathbf{x}, \mathbf{z}, V_\tau(\mathbf{x}, \mathbf{z})).$$

Suppose that $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ and $\boldsymbol{\mu} \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ are given. For any timestep $\tau > 0$ we would like to minimize the acceleration $A_\tau(\boldsymbol{\mu}, \boldsymbol{\gamma})$ over all $\boldsymbol{\gamma} \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ with

- (1) the transport plan taking $\mathbb{P}^1 \# \boldsymbol{\mu}$ to $\mathbb{P}^1 \# \boldsymbol{\gamma}$ is monotone,
- (2) the velocity distribution of $\boldsymbol{\gamma}$ is tangent to C_ϱ at the new configuration

(where the tangency to C_ϱ is yet to be specified). As mentioned above, this would be consistent with the usual setting in continuum mechanics. Unfortunately, tangent cones often do not possess good continuity properties, as can already be observed in convex polygons in \mathbf{R}^2 : the tangent cone at any point on an edge of the polygon is a half-space. But the tangent cone collapses to a smaller set at a corner. Consequently, the distance of a fixed point in \mathbf{R}^2 to the tangent cone may jump upwards as the base point of the tangent cone approaches a corner of the polygon.

We will therefore use an operator splitting: We first search for the transport that minimizes the acceleration cost, not imposing any restrictions on the final velocity, which will be determined a posteriori by formula (3.6). Then we project this new velocity onto the tangent cone (to be defined) at the new configuration. The second term in (3.5) now measures the cost of realizing a feasible velocity.

As explained above, if the velocity distribution of the second measure in (3.4) is not fixed, then the minimal acceleration cost simplifies. We therefore consider the following minimization problem: find the minimizer $\beta_\tau \in \mathcal{P}_2(\mathbf{R}^{3d})$ of

$$\beta \mapsto \frac{3}{4\tau^2} \int_{\mathbf{R}^{3d}} |(x + \tau\xi) - z|^2 \beta(d\mathbf{x}, dz)$$

among all $\beta \in \mathcal{P}_2(\mathbf{R}^{3d})$ with the following two properties:

$$(1.) \quad (\mathbb{P}^1, \mathbb{P}^2) \# \beta = \mu, \quad (2.) \quad (\mathbb{P}^1, \mathbb{P}^3) \# \beta \in C_\varrho. \quad (3.7)$$

It will be convenient to define $\mathbf{v}_\tau := (\mathbf{x}, \mathbf{x} + \tau\mathbf{v}) \# \mu$ and instead to minimize

$$\alpha \mapsto \frac{3}{4\tau^2} \int_{\mathbf{R}^{3d}} |y - z|^2 \alpha(dx, dy, dz)$$

over all $\alpha \in \mathcal{P}_2(\mathbf{R}^{3d})$ with the following two properties:

$$(1.) \quad (\mathbb{P}^1, \mathbb{P}^2) \# \alpha = \mathbf{v}_\tau, \quad (2.) \quad (\mathbb{P}^1, \mathbb{P}^3) \# \alpha \in C_\varrho. \quad (3.8)$$

Notice that for every $\tau > 0$ the push-forward under the map $(x, \xi) \mapsto (x + \tau\xi, \xi)$ with $(x, \xi) \in \mathbf{R}^{2d}$ is an automorphism between the spaces of measures $\alpha, \beta \in \mathcal{P}_2(\mathbf{R}^{3d})$ satisfying (3.8) and (3.7), respectively. We observe that (modulo the factor $3/4\tau^2$) we obtain exactly the minimization that defines the distance W_ϱ (see Theorem 2.4), where the second measure is allowed to range freely over the set C_ϱ . Therefore the minimization amounts to finding the element in C_ϱ closest to \mathbf{v}_τ with respect to the distance W_ϱ , i.e., to computing the metric projection onto C_ϱ .

3.3. Metric Projection. In order to study the minimization problem introduced in the previous section, we introduce on $\mathcal{P}_\varrho(\mathbf{R}^{2d})$ the analogues of scalar multiplication and vector addition in Hilbert spaces. This will allow us to define convexity of subsets of $\mathcal{P}_\varrho(\mathbf{R}^{2d})$ and metric projections onto such sets.

Definition 3.6 (Addition/Multiplication). Let $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ be given.

(i.) *Scaling.* For any $\gamma \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ and $s \in \mathbf{R}$ let

$$s\gamma := (\mathbf{x}, s\mathbf{y}^1) \# \gamma \in \mathcal{P}_\varrho(\mathbf{R}^{2d}).$$

(ii.) *Sum.* For any $\gamma^1, \gamma^2 \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ let

$$\gamma^1 \oplus \gamma^2 := \left\{ (\mathbf{x}, \mathbf{y}^1 + \mathbf{y}^2) \# \alpha : \alpha \in \text{ADM}_\varrho(\gamma^1, \gamma^2) \right\} \subset \mathcal{P}_\varrho(\mathbf{R}^{2d}).$$

If the plans are induced by functions, then the operations in Definition 3.6 reduce to the usual vector space structures on the Hilbert space $\mathcal{L}^2(\mathbf{R}^d, \varrho)$. Note also that for all $\gamma^1, \gamma^2 \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ and $s \in \mathbf{R}$ we have the useful equality

$$W_\varrho(s\gamma^1, s\gamma^2) = |s|W_\varrho(\gamma^1, \gamma^2).$$

We refer the reader to Section 4.1 in [41] for a proof.

Definition 3.7 (Closed Convex Cone). A nonempty subset $C \subset \mathcal{P}_\varrho(\mathbf{R}^{2d})$ will be called a closed convex set if it has the following two properties:

- (i.) **Closed.** Consider $\gamma^k \in C$ and $\gamma \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ with

$$W_\varrho(\gamma^k, \gamma) \longrightarrow 0 \quad \text{as } k \rightarrow \infty.$$

Then also $\gamma \in C$.

- (ii.) **Convex.** For all $\gamma^1, \gamma^2 \in C$ and $s \in [0, 1]$ we have

$$(1-s)\gamma^1 \oplus s\gamma^2 \in C. \quad (3.9)$$

The set C is a closed convex cone if it also has the following property:

- (iii.) **Cone.** For all $\gamma \in C$ and $s \geq 0$ we have $s\gamma \in C$.

We consider metric projections onto closed convex sets in $\mathcal{P}_\varrho(\mathbf{R}^{2d})$. They have similar properties like projections in Hilbert spaces.

Proposition 3.8 (Metric Projection). *Let $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ be given and $C \subset \mathcal{P}_\varrho(\mathbf{R}^{2d})$ a closed convex set. For any $\mathbf{v} \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ there is a unique $\mathbb{P}_C(\mathbf{v}) \in C$ with*

$$W_\varrho(\mathbf{v}, \mathbb{P}_C(\mathbf{v})) \leq W_\varrho(\mathbf{v}, \boldsymbol{\eta}) \quad \text{for all } \boldsymbol{\eta} \in C.$$

For every $\boldsymbol{\eta} \in C$ and all $\beta \in \mathcal{P}_2(\mathbf{R}^{4d})$ with

$$\begin{aligned} (\mathbb{X}, \mathbb{Y}^1, \mathbb{Y}^2) \# \beta &\in \text{OPT}_\varrho(\mathbf{v}, \mathbb{P}_C(\mathbf{v})), \\ (\mathbb{X}, \mathbb{Y}^1, \mathbb{Y}^3) \# \beta &\in \text{ADM}_\varrho(\mathbf{v}, \boldsymbol{\eta}). \end{aligned} \quad (3.10)$$

we have the inequality

$$\int_{\mathbf{R}^{4d}} \langle y^1 - y^2, y^2 - y^3 \rangle \beta(dx, dy^1, dy^2, dy^3) \geq 0. \quad (3.11)$$

Conversely, assume that there exists a $\zeta \in \mathcal{C}$ with the following property: for all $\boldsymbol{\eta} \in C$ there exists $\beta \in \mathcal{P}_2(\mathbf{R}^{4d})$ with

$$\begin{aligned} (\mathbb{X}, \mathbb{Y}^1, \mathbb{Y}^2) \# \beta &\in \text{ADM}_\varrho(\mathbf{v}, \zeta), \\ (\mathbb{X}, \mathbb{Y}^1, \mathbb{Y}^3) \# \beta &\in \text{OPT}_\varrho(\mathbf{v}, \boldsymbol{\eta}), \end{aligned} \quad (3.12)$$

such that inequality (3.11) is true. Then $\zeta = \mathbb{P}_C(\mathbf{v})$.

For any $\mathbf{v}^1, \mathbf{v}^2 \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ and any $\omega \in \mathcal{P}_2(\mathbf{R}^{5d})$ such that

$$\begin{aligned} (\mathbb{X}, \mathbb{Y}^1, \mathbb{Y}^2) \# \omega &\in \text{OPT}_\varrho(\mathbf{v}^1, \mathbb{P}_C(\mathbf{v}^1)), \\ (\mathbb{X}, \mathbb{Y}^3, \mathbb{Y}^4) \# \omega &\in \text{OPT}_\varrho(\mathbf{v}^2, \mathbb{P}_C(\mathbf{v}^2)), \end{aligned} \quad (3.13)$$

we can estimate as follows:

$$\int_{\mathbf{R}^{5d}} |y^2 - y^4|^2 \omega(dx, dy^1, \dots, dy^4) \leq \int_{\mathbf{R}^{5d}} |y^1 - y^3|^2 \omega(dx, dy^1, \dots, dy^4). \quad (3.14)$$

In particular, we have the contraction $W_\varrho(\mathbb{P}_C(\mathbf{v}^1), \mathbb{P}_C(\mathbf{v}^2)) \leq W_\varrho(\mathbf{v}^1, \mathbf{v}^2)$.

Proof. The proof is similar to the one of Proposition 4.30 in [41].

Step 1. Let $d := \inf\{W_\varrho(\mathbf{v}, \boldsymbol{\eta}) : \boldsymbol{\eta} \in C\} \geq 0$ and consider a sequence of plans $\boldsymbol{\eta}^n \in C$ such that $W_\varrho(\mathbf{v}, \boldsymbol{\eta}^n) \rightarrow d$ as $n \rightarrow \infty$. For any pair of indices $m, n \in \mathbf{N}$ choose $\beta^{m,n} \in \mathcal{P}_2(\mathbf{R}^{4d})$ with the property that

$$\begin{aligned} (\mathbb{x}, \mathbb{y}^1, \mathbb{y}^2) \# \beta^{m,n} &\in \text{OPT}_\varrho(\mathbf{v}, \boldsymbol{\eta}^m), \\ (\mathbb{x}, \mathbb{y}^1, \mathbb{y}^3) \# \beta^{m,n} &\in \text{OPT}_\varrho(\mathbf{v}, \boldsymbol{\eta}^n), \end{aligned}$$

and define the plans

$$\begin{aligned} \boldsymbol{\alpha}^{m,n} &:= (\mathbb{x}, \mathbb{y}^2, \mathbb{y}^3) \# \beta^{m,n} \in \text{ADM}_\varrho(\boldsymbol{\eta}^m, \boldsymbol{\eta}^n), \\ \boldsymbol{\eta}^{m,n} &:= (\mathbb{x}, \tfrac{1}{2}\mathbb{y}^2 + \tfrac{1}{2}\mathbb{y}^3) \# \beta^{m,n} \in C. \end{aligned}$$

The last inclusion follows from convexity (3.9). We claim that the sequence $\{\boldsymbol{\eta}^n\}_n$ is a Cauchy sequence with respect to W_ϱ . Indeed we have

$$\begin{aligned} &\frac{1}{2}W_\varrho(\boldsymbol{\eta}^m, \boldsymbol{\eta}^n)^2 \\ &\leq \int_{\mathbf{R}^{4d}} \frac{1}{2}|y^2 - y^3|^2 \beta^{m,n}(dx, dy^1 \dots dy^3) \\ &= \int_{\mathbf{R}^{4d}} \left(|y^1 - y^2|^2 + |y^1 - y^3|^2 - 2 \left| y^1 - \frac{y^2 + y^3}{2} \right|^2 \right) \beta^{m,n}(dx, dy^1 \dots dy^3) \\ &\leq W_\varrho(\mathbf{v}, \boldsymbol{\eta}^m)^2 + W_\varrho(\mathbf{v}, \boldsymbol{\eta}^n)^2 - 2W_\varrho(\mathbf{v}, \boldsymbol{\eta}^{m,n})^2 \end{aligned}$$

for all $m, n \in \mathbf{N}$. Notice that $W_\varrho(\mathbf{v}, \boldsymbol{\eta}^{m,n}) \geq d$ because $\boldsymbol{\eta}^{m,n} \in C$. This yields

$$\frac{1}{2}W_\varrho(\boldsymbol{\eta}^m, \boldsymbol{\eta}^n)^2 \leq W_\varrho(\mathbf{v}, \boldsymbol{\eta}^m)^2 + W_\varrho(\mathbf{v}, \boldsymbol{\eta}^n)^2 - 2d^2. \quad (3.15)$$

Since by assumption the sequence $\{\boldsymbol{\eta}^n\}_n$ is minimizing, the right-hand side of (3.15) converges to zero as $m, n \rightarrow \infty$, which proves our claim. Recall that $(\mathcal{P}_\varrho(\mathbf{R}^{2d}), W_\varrho)$ is a complete metric space. It follows that there is a $\mathbb{P}_C(\mathbf{v}) \in C$ with the property that $W_\varrho(\boldsymbol{\eta}^n, \mathbb{P}_C(\mathbf{v})) \rightarrow 0$. By lower semicontinuity of the distance W_ϱ , we now have $W_\varrho(\mathbf{v}, \mathbb{P}_C(\mathbf{v})) = d$. This establishes the existence of a minimizer.

Step 2. To prove uniqueness, assume that there exists $\boldsymbol{\eta} \in C$ with $W_\varrho(\mathbf{v}, \boldsymbol{\eta}) = d$. Now choose a plan $\beta \in \mathcal{P}_2(\mathbf{R}^{4d})$ that satisfies

$$\begin{aligned} (\mathbb{x}, \mathbb{y}^1, \mathbb{y}^2) \# \beta &\in \text{OPT}_\varrho(\mathbf{v}, \mathbb{P}_C(\mathbf{v})), \\ (\mathbb{x}, \mathbb{y}^1, \mathbb{y}^3) \# \beta &\in \text{OPT}_\varrho(\mathbf{v}, \boldsymbol{\eta}), \end{aligned}$$

and define the plans

$$\begin{aligned} \boldsymbol{\alpha} &:= (\mathbb{x}, \mathbb{y}^1, \tfrac{1}{2}\mathbb{y}^2 + \tfrac{1}{2}\mathbb{y}^3) \# \beta \in \text{ADM}_\varrho(\mathbf{v}, \bar{\boldsymbol{\eta}}), \\ \bar{\boldsymbol{\eta}} &:= (\mathbb{x}, \tfrac{1}{2}\mathbb{y}^2 + \tfrac{1}{2}\mathbb{y}^3) \# \beta \in C. \end{aligned}$$

The last inclusion again follows from convexity (3.9). We can then estimate

$$\begin{aligned}
 2W_\varrho(\mathbf{v}, \bar{\boldsymbol{\eta}})^2 &\leq \int_{\mathbf{R}^{4d}} 2 \left| y^1 - \frac{y^2 + y^3}{2} \right|^2 \beta(dx, dy^1 \dots dy^3) \\
 &= \int_{\mathbf{R}^{4d}} \left(|y^1 - y^2|^2 + |y^1 - y^3|^2 - \frac{1}{2}|y^2 - y^3|^2 \right) \beta(dx, dy^1 \dots dy^3) \\
 &\leq W_\varrho(\mathbf{v}, \mathbb{P}_C(\mathbf{v}))^2 + W_\varrho(\mathbf{v}, \boldsymbol{\eta})^2 - \frac{1}{2} W_\varrho(\mathbb{P}_C(\mathbf{v}), \boldsymbol{\eta})^2.
 \end{aligned} \tag{3.16}$$

By our choice of $\boldsymbol{\eta}$, we obtain $W_\varrho(\mathbf{v}, \bar{\boldsymbol{\eta}})^2 \leq d^2 - \frac{1}{4} W_\varrho(\mathbb{P}_C(\mathbf{v}), \boldsymbol{\eta})^2$, which shows that if $\mathbb{P}_C(\mathbf{v})$ and $\boldsymbol{\eta}$ are different, then $W_\varrho(\mathbf{v}, \bar{\boldsymbol{\eta}}) < d$. This contradicts the definition of d because $\bar{\boldsymbol{\eta}} \in C$. Therefore the minimizer must be unique.

Step 3. For any $\boldsymbol{\eta} \in C$ consider now $\beta \in \mathcal{P}_2(\mathbf{R}^{4d})$ with (3.10). For every $s > 0$ we define $\boldsymbol{\eta}_s := (\mathbb{x}, (1-s)\mathbb{y}^2 + s\mathbb{y}^3) \# \beta \in C$; see (3.9). Then

$$\begin{aligned}
 \int_{\mathbf{R}^{4d}} |y^1 - y^2|^2 \beta(dx, dy^1, dy^2, dy^3) &= W_\varrho(\mathbf{v}, \mathbb{P}_C(\mathbf{v}))^2 \\
 &\leq W_\varrho(\mathbf{v}, \boldsymbol{\eta}_s)^2 \leq \int_{\mathbf{R}^{4d}} |y^1 - ((1-s)y^2 + sy^3)|^2 \beta(dx, dy^1, dy^2, dy^3),
 \end{aligned}$$

which implies the estimate

$$\begin{aligned}
 0 &\geq -2s \int_{\mathbf{R}^{4d}} \langle y^1 - y^2, y^2 - y^3 \rangle \beta(dx, dy^1, dy^2, dy^3) \\
 &\quad - s^2 \int_{\mathbf{R}^{4d}} |y^2 - y^3|^2 \beta(dx, dy^1, dy^2, dy^3).
 \end{aligned} \tag{3.17}$$

Notice that the second integral on the right-hand side of (3.17) is finite. Dividing the inequality (3.17) by $-2s < 0$ and letting $s \rightarrow 0$, we obtain (3.11).

Conversely, let $\boldsymbol{\zeta} \in C$. Assume that for every $\boldsymbol{\eta} \in C$ there exists $\beta \in \mathcal{P}_2(\mathbf{R}^{4d})$ with (3.12) satisfying (3.11). Then we can estimate as follows:

$$\begin{aligned}
 W_\varrho(\mathbf{v}, \boldsymbol{\zeta})^2 - W_\varrho(\mathbf{v}, \boldsymbol{\eta})^2 &\leq \int_{\mathbf{R}^{4d}} \left(|y^1 - y^2|^2 - |y^1 - y^3|^2 \right) \beta(dx, dy^1, dy^2, dy^3) \\
 &= -2 \int_{\mathbf{R}^{4d}} \langle y^1 - y^2, y^2 - y^3 \rangle \beta(dx, dy^1, dy^2, dy^3) \\
 &\quad - \int_{\mathbf{R}^{4d}} |y^2 - y^3|^2 \beta(dx, dy^1, dy^2, dy^3),
 \end{aligned}$$

which is nonpositive, by assumption. Since $\boldsymbol{\eta} \in C$ was arbitrary, the plan $\boldsymbol{\zeta}$ must be equal to the uniquely determined metric projection $\mathbb{P}_C(\mathbf{v})$.

Step 4. Consider now $\mathbf{v}^1, \mathbf{v}^2 \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ and their metric projections onto C . For all $\boldsymbol{\alpha} \in \text{ADM}_\varrho(\mathbf{v}^1, \mathbf{v}^2)$ there exists $\boldsymbol{\omega} \in \mathcal{P}_2(\mathbf{R}^{5d})$ with $(\mathbb{x}, \mathbb{y}^1, \mathbb{y}^3) \# \boldsymbol{\omega} = \boldsymbol{\alpha}$ and (3.13). Since $(\mathbb{x}, \mathbb{y}^1, \mathbb{y}^4) \# \boldsymbol{\omega} \in \text{ADM}_\varrho(\mathbf{v}^1, \mathbb{P}_C(\mathbf{v}^2))$, we apply (3.11) and obtain

$$\int_{\mathbf{R}^{5d}} \langle y^1 - y^2, y^2 - y^4 \rangle \boldsymbol{\omega}(dx, dy^1, \dots, dy^4) \geq 0. \tag{3.18}$$

Similarly, since $(\mathbb{x}, \mathbb{y}^3, \mathbb{y}^2) \# \boldsymbol{\omega} \in \text{ADM}_\varrho(\mathbf{v}^2, \mathbb{P}_C(\mathbf{v}^1))$, we have

$$\int_{\mathbf{R}^{5d}} \langle y^3 - y^4, y^4 - y^2 \rangle \boldsymbol{\omega}(dx, dy^1, \dots, dy^4) \geq 0. \tag{3.19}$$

Adding (3.18) and (3.19) and using the Cauchy-Schwarz inequality, we get (3.14). The left-hand side of (3.14) is always bigger than or equal to $W_\varrho(\mathbb{P}_C(\mathbf{v}^1), \mathbb{P}_C(\mathbf{v}^2))^2$. The right-hand side equals $W_\varrho(\mathbf{v}^1, \mathbf{v}^2)^2$ whenever $\alpha \in \text{OPT}_\varrho(\mathbf{v}^1, \mathbf{v}^2)$. \square

3.4. Non-Splitting Projections. Under a suitable assumption on the closed convex set, the projections in Proposition 3.8 can be expressed in terms of maps.

Assumption 3.9. For any $\eta \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ and $\zeta \in C$ we consider the disintegrations $\eta(dx, dy) =: \eta_x(dy) \varrho(dx)$ and $\zeta(dx, dy) =: \zeta_x(dy) \varrho(dx)$. We assume that if

$$\text{spt } \eta_x \subset \overline{\text{conv}}(\text{spt } \zeta_x) \quad \text{for } \varrho\text{-a.e. } x \in \mathbf{R}^d, \quad (3.20)$$

then also $\eta \in C$.

Proposition 3.10 (Properties of $\mathbb{P}_C(\mathbf{v})$). *Let the closed convex set $C \subset \mathcal{P}_\varrho(\mathbf{R}^{2d})$ satisfy Assumption 3.9 and let $\mathbb{P}_C(\mathbf{v})$ be the metric projection of $\mathbf{v} \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ onto C ; see Proposition 3.8. Then there exists a unique $\mathbb{Z}_\mathbf{v} \in \mathcal{L}^2(\mathbf{R}^{2d}, \mathbf{v})$ with*

$$\text{OPT}_\varrho(\mathbf{v}, \mathbb{P}_C(\mathbf{v})) = \{(\mathbb{X}, \mathbb{Y}, \mathbb{Z}_\mathbf{v}) \# \mathbf{v}\}. \quad (3.21)$$

Proof. The proof is similar to the one of Propositions 4.32 in [41].

Step 1. Fix any $\alpha \in \text{OPT}_\varrho(\mathbf{v}, \mathbb{P}_C(\mathbf{v}))$ and consider the disintegration

$$\alpha(dx, dy, dz) =: \alpha_{(x,y)}(dz) \mathbf{v}(dx, dy). \quad (3.22)$$

Then we define the function

$$\mathbb{Z}_\mathbf{v}(x, y) := \int_{\mathbf{R}^d} z \alpha_{(x,y)}(dz) \quad \text{for } \mathbf{v}\text{-a.e. } (x, y) \in \mathbf{R}^{2d}, \quad (3.23)$$

and the plans

$$\begin{aligned} \bar{\alpha} &:= (\mathbb{X}, \mathbb{Y}, \mathbb{Z}_\mathbf{v}) \# \mathbf{v} \in \text{ADM}_\varrho(\mathbf{v}, \bar{\mathbf{v}}), \\ \bar{\mathbf{v}} &:= (\mathbb{X}, \mathbb{Z}_\mathbf{v}) \# \mathbf{v}. \end{aligned}$$

We claim that $\bar{\mathbf{v}} \in C$. Notice first that clearly

$$\mathbb{Z}_\mathbf{v}(x, y) \subset \overline{\text{conv}}(\text{spt } \alpha_{(x,y)}) \quad \text{for } \mathbf{v}\text{-a.e. } (x, y) \in \mathbf{R}^{2d}.$$

Now consider the disintegrations

$$\begin{aligned} \bar{\mathbf{v}}(dx, dz) &=: \bar{v}_x(dz) \varrho(dx), \\ \mathbf{v}(dx, dy) &=: v_x(dy) \varrho(dx), \\ \mathbb{P}_C(\mathbf{v})(dx, dz) &=: \mathbb{P}_C(\mathbf{v})_x(dz) \varrho(dx). \end{aligned}$$

It follows that

$$\mathbb{P}_C(\mathbf{v})_x(dz) = \int_{\mathbf{R}^d} \alpha_{(x,y)}(dz) v_x(dy) \quad \text{for } \varrho\text{-a.e. } x \in \mathbf{R}^d,$$

and so $\text{spt } \alpha_{(x,y)} \subset \text{spt } \mathbb{P}_C(\mathbf{v})_x$ for \mathbf{v}_x -a.e. $y \in \mathbf{R}^d$. This yields

$$\text{spt } \bar{v}_x \subset \overline{\text{conv}}(\text{spt } \mathbb{P}_C(\mathbf{v})_x) \quad \text{for } \varrho\text{-a.e. } x \in \mathbf{R}^d.$$

By Assumption 3.9, this implies that $\bar{\mathbf{v}} \in C$.

Using that $\bar{\alpha} \in \text{ADM}_\ell(\mathbf{v}, \bar{\mathbf{v}})$, we now estimate

$$\begin{aligned} W_\ell(\mathbf{v}, \bar{\mathbf{v}})^2 &\leq \int_{\mathbf{R}^{2d}} |y - \mathbb{Z}_\mathbf{v}(x, y)|^2 \mathbf{v}(dx, dy) \\ &= \int_{\mathbf{R}^{2d}} \left| \int_{\mathbf{R}^d} (y - z) \alpha_{(x, y)}(dz) \right|^2 \mathbf{v}(dx, dy) \\ &\leq \int_{\mathbf{R}^{2d}} \int_{\mathbf{R}^d} |y - z|^2 \alpha_{(x, y)}(dz) \mathbf{v}(dx, dy) = W_\ell(\mathbf{v}, \mathbb{P}_C(\mathbf{v}))^2. \end{aligned}$$

The first equality follows from (3.23) and the second one from (3.22). For the second inequality we have used Jensen's inequality. Recall that Jensen's inequality is strict unless the probability measure is a Dirac measure, which implies that if $\alpha_{(x, y)}$ is not a Dirac measure for \mathbf{v} -a.e. $(x, y) \in \mathbf{R}^{2d}$, then $W_\ell(\mathbf{v}, \bar{\mathbf{v}}) < W_\ell(\mathbf{v}, \mathbb{P}_C(\mathbf{v}))$. This contradicts the definition of $\mathbb{P}_C(\mathbf{v})$ since $\bar{\mathbf{v}} \in C$. We conclude that

$$\alpha_{(x, y)}(dz) = \delta_{\mathbb{Z}_\mathbf{v}(x, y)}(dz) \quad \text{for } \mathbf{v}\text{-a.e. } (x, y) \in \mathbf{R}^{2d},$$

and thus $\alpha = \bar{\alpha}$. The same argument works for all $\alpha \in \text{OPT}_\ell(\mathbf{v}, \mathbb{P}_C(\mathbf{v}))$, and so all optimal transport plans between \mathbf{v} and $\mathbb{P}_C(\mathbf{v})$ are induced by maps.

Step 2. To prove uniqueness, assume there exist two maps $\mathbb{Z}^1, \mathbb{Z}^2 \in \mathcal{L}^2(\mathbf{R}^{2d}, \mathbf{v})$ such that $\alpha^k := (\mathbb{x}, \mathbb{y}, \mathbb{Z}^k) \# \mathbf{v} \in \text{OPT}_\ell(\mathbf{v}, \mathbb{P}_C(\mathbf{v}))$ for $k = 1..2$. Let

$$\begin{aligned} \bar{\beta} &:= (\mathbb{x}, \mathbb{y}, \mathbb{Z}^1, \mathbb{Z}^2) \# \mathbf{v}, \\ \bar{\alpha} &:= (\mathbb{x}, \mathbb{y}, \tfrac{1}{2}\mathbb{y}^2 + \tfrac{1}{2}\mathbb{y}^3) \# \bar{\beta}. \end{aligned}$$

We claim that $\bar{\alpha} \in \text{OPT}_\ell(\mathbf{v}, \mathbb{P}_C(\mathbf{v}))$. If this is true, and if \mathbb{Z}^1 and \mathbb{Z}^2 are different, then $\bar{\alpha}$ is not induced by a map, in contradiction to what we proved in Step 2. We can therefore define $\mathbb{Z}_\mathbf{v}$ unambiguously by the property (3.21).

To prove the claim, let $\bar{\mathbf{v}} := (\mathbb{x}, \mathbb{y}^2) \# \bar{\alpha} = (\mathbb{x}, \tfrac{1}{2}\mathbb{y}^2 + \tfrac{1}{2}\mathbb{y}^3) \# \bar{\beta}$ and note that

$$(\mathbb{x}, \mathbb{y}^2, \mathbb{y}^3) \# \bar{\beta} \in \text{ADM}_\ell(\mathbb{P}_C(\mathbf{v}), \mathbb{P}_C(\mathbf{v})).$$

Then $\bar{\mathbf{v}} \in C$ because of convexity (3.9). We have $\bar{\alpha} \in \text{ADM}_\ell(\mathbf{v}, \bar{\mathbf{v}})$ and

$$\begin{aligned} (\mathbb{x}, \mathbb{y}^1, \mathbb{y}^2) \# \bar{\beta} &\in \text{OPT}_\ell(\mathbf{v}, \mathbb{P}_C(\mathbf{v})), \\ (\mathbb{x}, \mathbb{y}^1, \mathbb{y}^3) \# \bar{\beta} &\in \text{OPT}_\ell(\mathbf{v}, \mathbb{P}_C(\mathbf{v})). \end{aligned}$$

Arguing as in estimate (3.16), we obtain that $W_\ell(\mathbf{v}, \bar{\mathbf{v}}) \leq W_\ell(\mathbf{v}, \mathbb{P}_C(\mathbf{v})) = d$, which shows that $\bar{\mathbf{v}} = \mathbb{P}_C(\mathbf{v})$, by uniqueness of the minimizer. \square

Remark 3.11. Under Assumption 3.9, the third part of Proposition 3.8 simplifies as follows: for any plans $\mathbf{v}^1, \mathbf{v}^2 \in \mathcal{P}_\ell(\mathbf{R}^{2d})$ let $\mathbb{Z}^k \in \mathcal{L}^2(\mathbf{R}^{2d}, \mathbf{v}^k)$ be defined as in (3.21) for $k = 1..2$. Then we have the following inequality:

$$\begin{aligned} &\int_{\mathbf{R}^{3d}} |\mathbb{Z}^1(x, y^1) - \mathbb{Z}^2(x, y^2)|^2 \alpha(dx, dy^1, dy^2) \\ &\leq \int_{\mathbf{R}^{3d}} |y^1 - y^2|^2 \alpha(dx, dy^1, dy^2) \quad \text{for all } \alpha \in \text{ADM}_\ell(\mathbf{v}^1, \mathbf{v}^2). \end{aligned}$$

Remark 3.12. If C is a closed convex cone, then (3.11) implies the following statement: for every $\boldsymbol{\eta} \in C$ and all $\boldsymbol{\alpha} \in \text{ADM}_\varrho(\boldsymbol{v}, \boldsymbol{\eta})$ we have that

$$\int_{\mathbf{R}^{3d}} \langle y - \mathbb{Z}_{\boldsymbol{v}}(x, y), \mathbb{Z}_{\boldsymbol{v}}(x, y) \rangle \boldsymbol{v}(dx, dy) = 0, \quad (3.24)$$

$$\int_{\mathbf{R}^{3d}} \langle y - \mathbb{Z}_{\boldsymbol{v}}(x, y), z \rangle \boldsymbol{\alpha}(dx, dy, dz) \leq 0. \quad (3.25)$$

Indeed note first that because of (3.21), the inequality (3.11) reads as follows:

$$\int_{\mathbf{R}^{3d}} \langle y - \mathbb{Z}_{\boldsymbol{v}}(x, y), \mathbb{Z}_{\boldsymbol{v}}(x, y) - z \rangle \boldsymbol{\alpha}(dx, dy, dz) \geq 0 \quad (3.26)$$

for all $\boldsymbol{\eta}, \boldsymbol{\alpha}$ as above. We have $\boldsymbol{\eta}_0 := (\text{id}, 0) \# \varrho \in C$ since C is a cone. Then

$$\boldsymbol{\alpha}^1 := (\mathbf{x}, \mathbf{y}, 0) \# \boldsymbol{v} \in \text{ADM}_\varrho(\boldsymbol{v}, \boldsymbol{\eta}_0).$$

Using $\boldsymbol{\alpha}^1$ in (3.26), we obtain the inequality

$$\int_{\mathbf{R}^{2d}} \langle y - \mathbb{Z}_{\boldsymbol{v}}(x, y), \mathbb{Z}_{\boldsymbol{v}}(x, y) \rangle \boldsymbol{v}(dx, dy) \geq 0. \quad (3.27)$$

On the other hand, we have $2\mathbb{P}_C(\boldsymbol{v}) \in C$ since C is a cone. Then

$$\boldsymbol{\alpha}^2 := (\mathbf{x}, \mathbf{y}, 2\mathbb{Z}_{\boldsymbol{v}}) \# \boldsymbol{v} \in \text{ADM}_\varrho(\boldsymbol{v}, 2\mathbb{P}_C(\boldsymbol{v})).$$

Using $\boldsymbol{\alpha}^2$ in (3.26), we obtain the inequality

$$\int_{\mathbf{R}^{2d}} \langle y - \mathbb{Z}_{\boldsymbol{v}}(x, y), -\mathbb{Z}_{\boldsymbol{v}}(x, y) \rangle \boldsymbol{v}(dx, dy) \geq 0. \quad (3.28)$$

We now combine (3.27) and (3.28), and get (3.24) and thus (3.25).

3.5. Monotone Transport Plans. Propositions 3.8 and 3.10 apply to C_ϱ .

Proposition 3.13 (Monotone Transport Plans). *Let $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$. Then*

$$C_\varrho := \left\{ \boldsymbol{\gamma} \in \mathcal{P}_\varrho(\mathbf{R}^{2d}) : \text{spt } \boldsymbol{\gamma} \text{ is a monotone subset of } \mathbf{R}^d \times \mathbf{R}^d \right\}$$

(which is (3.2)) is a closed convex cone and Assumption 3.9 is satisfied.

Proof. We proceed in three steps.

Step 1. Consider first plans $\boldsymbol{\gamma}^k$ and $\boldsymbol{\gamma}$ as in Definition 3.7 (i.). Since

$$W(\boldsymbol{\gamma}^k, \boldsymbol{\gamma}) \leq W_\varrho(\boldsymbol{\gamma}^k, \boldsymbol{\gamma})$$

we have that $\boldsymbol{\gamma}^k \rightarrow \boldsymbol{\gamma}$ with respect to the Wasserstein distance, and thus narrowly; see Proposition 7.1.5 in [4]. One can check that $\boldsymbol{\gamma} \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$. Fix $(x_i, y_i) \in \text{spt } \boldsymbol{\gamma}$ with $i = 1..2$. Since narrow convergence of probability measures implies Kuratowski convergence of their supports (see Proposition 5.1.8 in [4]), there exist

$$(x_i^k, y_i^k) \in \text{spt } \boldsymbol{\gamma}^k \quad \text{such that} \quad \lim_{k \rightarrow \infty} (x_i^k, y_i^k) = (x_i, y_i)$$

for $i = 1..2$. Since $\text{spt } \boldsymbol{\gamma}^k$ is monotone for all k , we obtain that

$$\langle x_1 - x_2, y_1 - y_2 \rangle = \lim_{k \rightarrow \infty} \langle x_1^k - x_2^k, y_1^k - y_2^k \rangle \geq 0.$$

Since the (x^i, y^i) were arbitrary, we conclude that $\boldsymbol{\gamma} \in C_\varrho$.

Step 2. Let now $\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2 \in C_\varrho$ and $s \in [0, 1]$ be given. For any $\boldsymbol{\alpha} \in \text{ADM}_\varrho(\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2)$ we define the interpolation transport plan $\boldsymbol{\gamma}_s := (\mathbf{x}, (1-s)\mathbf{y}^1 + s\mathbf{y}^2) \# \boldsymbol{\alpha} \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$.

Consider now any point $(x, y) \in \text{spt } \gamma_s$. By the definition of support of a measure, for all $\varepsilon > 0$ there exists $(\hat{x}, \hat{y}^1, \hat{y}^2) \in \text{spt } \alpha$ such that

$$\hat{x} \in B_\varepsilon(x) \quad \text{and} \quad (1-s)\hat{y}^1 + s\hat{y}^2 \in B_\varepsilon(y). \quad (3.29)$$

Indeed, assume there exists an $\varepsilon > 0$ with the property that for all $(\hat{x}, \hat{y}^1, \hat{y}^2) \in \text{spt } \alpha$ statement (3.29) is wrong. Then $\gamma_s(B_\varepsilon(x) \times B_\varepsilon(y)) = 0$, which is a contradiction to our choice $(x, y) \in \text{spt } \gamma_s$. Now $(\hat{x}, \hat{y}^1, \hat{y}^2) \in \text{spt } \alpha$ implies that

$$0 < \alpha(B_r(\hat{x}) \times B_r(\hat{y}^1) \times B_r(\hat{y}^2)) \leq \gamma^k(B_r(\hat{x}) \times B_r(\hat{y}^k))$$

for all $r > 0$, with $k = 1..2$. We conclude that $(\hat{x}, \hat{y}^k) \in \text{spt } \gamma^k$.

We can now apply the above argument to a pair of points $(x_i, y_i) \in \text{spt } \gamma_s$, with $i = 1..2$. For any $\varepsilon > 0$ we find $(\hat{x}_i, \hat{y}_i^k) \in \text{spt } \gamma^k$, $i = 1..2$, such that

$$\hat{x}_i \in B_\varepsilon(x_i) \quad \text{and} \quad (1-s)\hat{y}_i^1 + s\hat{y}_i^2 \in B_\varepsilon(y_i).$$

Since $\text{spt } \gamma^k$ is monotone, we obtain the estimate

$$\begin{aligned} \langle x_1 - x_2, y_1 - y_2 \rangle &\geq (1-s)\langle \hat{x}_1 - \hat{x}_2, \hat{y}_1^1 - \hat{y}_2^1 \rangle + s\langle \hat{x}_1 - \hat{x}_2, \hat{y}_1^2 - \hat{y}_2^2 \rangle - 4(M + \varepsilon)\varepsilon \\ &\geq -4(M + \varepsilon)\varepsilon, \end{aligned}$$

with $M := \max_i \{|x_i|, |y_i|\}$. Since $\varepsilon > 0$ and $(x_i, y_i) \in \text{spt } \gamma_s$ were arbitrary, we get that $\text{spt } \gamma_s$ is monotone. Since $\alpha \in \text{ADM}_\varrho(\gamma^1, \gamma^2)$ was arbitrary, we obtain (3.9). In a similar way, one proves that if $\gamma \in C_\varrho$, then also $s\gamma \in C_\varrho$ for all $s \geq 0$.

Step 3. In order to prove Assumption 3.9, note that if $\zeta \in C_\varrho$, then its support is contained in the graph of a maximal monotone set-valued map u (we may consider a suitable extension if necessary). For ϱ -a.e. $x \in \mathbf{R}^d$ we have $\text{spt } \zeta_x \subset u(x)$, which is a closed and convex set; see [1]. Then $\text{spt } \eta_x \subset u(x)$ as well because of assumption (3.20), which implies that the support of η is monotone and hence $\eta \in C_\varrho$. \square

Remark 3.14. Proposition 3.10 implies that whenever $\mathbf{v} \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ is induced by a map, i.e., there exists a $\mathcal{Y} \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$ taking values in \mathbf{R}^d such that

$$\mathbf{v}(dx, dy) = \delta_{\mathcal{Y}(x)}(dy) \varrho(dx),$$

then the projection $\mathbb{P}_{C_\varrho}(\mathbf{v})$ is induced by a map as well:

$$\mathbb{P}_{C_\varrho}(\mathbf{v})(dx, dz) = \delta_{\mathbb{Z}_\mathbf{v}(x, \mathcal{Y}(x))}(dz) \varrho(dx).$$

Notice that if ϱ is absolutely continuous with respect to the Lebesgue measure, then all monotone transport plans in C_ϱ are in fact induced by maps. This follows in the same way as for optimal transport plans (which are contained in the subdifferentials of convex functions, thus monotone): the set of points where a (maximal) monotone set-valued map is multi-valued is a Lebesgue null set; see [1].

4. ENERGY MINIMIZATION: PRESSURELESS GASES

For our variational time discretization of the pressureless gas dynamics equations (1.3), we divide the time interval $[0, T]$ into subintervals of length $\tau > 0$. For every timestep, we minimize the acceleration cost 3.5 over the cone of monotone transport plans. As explained in Section 3.2, this reduces to a metric projection, which further simplifies to a minimization over a closed convex cone in a Hilbert space: we may consider monotone transport *maps* instead of plans because of Proposition 3.10.

4.1. Configuration Manifold. Going back to our original setup, we will consider monotone transport maps that are defined on measures $\mu \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ representing the distribution of mass and *velocity*, not representing transport plans.

Definition 4.1 (Configurations). For any $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ and $\mu \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$, let

$$\mathcal{C}_\mu := \left\{ \mathbb{T} \in \mathcal{L}^2(\mathbf{R}^{2d}, \mu) : (\mathbb{x}, \mathbb{T}) \# \mu \in C_\varrho \right\}.$$

Lemma 4.2 (Closed Convex Cone). \mathcal{C}_μ is a closed convex cone in $\mathcal{L}^2(\mathbf{R}^{2d}, \mu)$.

Proof. We observe first that for any $\mathbb{T}^1, \mathbb{T}^2 \in \mathcal{C}_\mu$, we have that

$$(\mathbb{x}, \mathbb{T}^1, \mathbb{T}^2) \# \mu \in \text{ADM}_\varrho(\gamma^1, \gamma^2) \quad (4.1)$$

where $\gamma^n := (\mathbb{x}, \mathbb{T}^n) \# \mu \in C_\varrho$ with $n = 1..2$. This implies the estimate

$$W_\varrho(\gamma^1, \gamma^2) \leq \|\mathbb{T}^1 - \mathbb{T}^2\|_{\mathcal{L}^2(\mathbf{R}^{2d}, \mu)}. \quad (4.2)$$

Consider now a sequence $\mathbb{T}^k \rightarrow \mathbb{T}$ in $\mathcal{L}^2(\mathbf{R}^{2d}, \mu)$ and define $\gamma^k := (\mathbb{x}, \mathbb{T}^k) \# \mu$. Let $\gamma := (\mathbb{x}, \mathbb{T}) \# \mu$ and notice that $W_\varrho(\gamma^k, \gamma) \rightarrow 0$ because of (4.2). If now $\mathbb{T}^k \in \mathcal{C}_\mu$ and thus $\gamma^k \in C_\varrho$ for all k , then also $\gamma \in C_\varrho$ since C_ϱ is closed with respect to W_ϱ ; see Proposition 3.13. This proves that $\mathbb{T} \in \mathcal{C}_\mu$. For any $s \in [0, 1]$ we have

$$\gamma_s := (\mathbb{x}, (1-s)\mathbb{T}^1 + s\mathbb{T}^2) \# \mu = (\mathbb{x}, (1-s)\gamma^1 + s\gamma^2) \# \alpha^{1,2},$$

where $\alpha^{1,2} := (\mathbb{x}, \mathbb{T}^1, \mathbb{T}^2) \# \mu$ with $\mathbb{T}^n \in \mathcal{C}_\mu$ and $n = 1..2$. Using (4.1), we conclude that $\gamma_s \in (1-s)\gamma^1 \oplus s\gamma^2$ (see Definition 3.6), which is in C_ϱ , by Proposition 3.13. Hence $(1-s)\mathbb{T}^1 + s\mathbb{T}^2 \in \mathcal{C}_\mu$. The proof that \mathcal{C}_μ is a cone is analogous. \square

Remark 4.3. For every $\tau > 0$ let $\mathbb{Z}_{\mathbf{v}_\tau}$ be the unique map defined in Proposition 3.10 representing the metric projection of $\mathbf{v}_\tau := (\mathbb{x}, \mathbb{x} + \tau\mathbf{v}) \# \mu$ onto the closed convex cone C_ϱ . We define $\mathbb{T}_\tau(x, \xi) := \mathbb{Z}_{\mathbf{v}_\tau}(x, x + \tau\xi)$ for μ -a.e. $(x, \xi) \in \mathbf{R}^{2d}$ so that

$$(\mathbb{x}, \mathbb{y}, \mathbb{Z}_{\mathbf{v}_\tau}) \# \mathbf{v}_\tau = (\mathbb{x}, \mathbb{x} + \tau\mathbf{v}, \mathbb{T}_\tau) \# \mu. \quad (4.3)$$

Then \mathbb{T}_τ must be the uniquely determined metric projection of $\mathbb{x} + \tau\mathbf{v} \in \mathcal{L}^2(\mathbf{R}^{2d}, \mu)$ onto the cone \mathcal{C}_μ (see [64] for more information about metric projections in Hilbert spaces). Indeed for any $\mathbb{S} \in \mathcal{C}_\mu$ (for which $\gamma_\mathbb{S} := (\mathbb{x}, \mathbb{S}) \# \mu \in C_\varrho$) we have

$$\begin{aligned} \|(\mathbb{x} + \tau\mathbf{v}) - \mathbb{T}_\tau\|_{\mathcal{L}^2(\mathbf{R}^{2d}, \mu)} &= \|\mathbb{y} - \mathbb{Z}_{\mathbf{v}_\tau}\|_{\mathcal{L}^2(\mathbf{R}^{2d}, \mathbf{v}_\tau)} \\ &= W_\varrho(\mathbf{v}_\tau, \mathbb{P}_{C_\varrho}(\mathbf{v}_\tau)) \\ &\leq W_\varrho(\mathbf{v}_\tau, \gamma_\mathbb{S}) \leq \|(\mathbb{x} + \tau\mathbf{v}) - \mathbb{S}\|_{\mathcal{L}^2(\mathbf{R}^{2d}, \mu)}. \end{aligned}$$

The first equality follows from definition (4.3) and the second one from (3.21). The subsequent inequality is true because $\mathbb{P}_{C_\varrho}(\mathbf{v}_\tau)$ is closest to \mathbf{v}_τ in C_ϱ with respect to W_ϱ , by definition. Finally, we have used that $(\mathbb{x}, \mathbb{x} + \tau\mathbf{v}, \mathbb{S}) \# \mu \in \text{ADM}_\varrho(\mathbf{v}_\tau, \gamma_\mathbb{S})$. We will write $\mathbb{T}_\tau = \mathbb{P}_{\mathcal{C}_\mu}(\mathbb{x} + \tau\mathbf{v})$. The map \mathbb{T}_τ is uniquely determined by

$$\int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathbb{T}_\tau(x, \xi), \mathbb{T}_\tau(x, \xi) \rangle \mu(dx, d\xi) = 0, \quad (4.4)$$

$$\int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathbb{T}_\tau(x, \xi), \mathbb{S}(x, \xi) \rangle \mu(dx, d\xi) \leq 0 \quad \text{for all } \mathbb{S} \in \mathcal{C}_\mu. \quad (4.5)$$

We just need to combine Lemma 1.1 in [64] with Remark 3.12.

Remark 4.4. A map $\mathbb{T} \in \mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu})$ is in \mathcal{C}_μ if and only if the following statement is true: There exists a Borel set $N_\mathbb{T} \subset \mathbf{R}^{2d}$ with $\boldsymbol{\mu}(N_\mathbb{T}) = 0$ such that

$$\langle \mathbb{T}(x_1, \xi_1) - \mathbb{T}(x_2, \xi_2), x_1 - x_2 \rangle \geq 0 \quad \text{for all } (x_i, \xi_i) \in \mathbf{R}^{2d} \setminus N_\mathbb{T} \quad (4.6)$$

with $i = 1..2$. Indeed consider any $\mathbb{T} \in \mathcal{C}_\mu$ and $\gamma_\mathbb{T} := (\mathbb{x}, \mathbb{T}) \# \boldsymbol{\mu} \in C_\varrho$. Then

$$\begin{aligned} N_\mathbb{T} &:= \left\{ (x, \xi) \in \mathbf{R}^{2d} : (x, \mathbb{T}(x, \xi)) \notin \text{spt } \gamma_\mathbb{T} \right\} \quad \text{satisfies} \\ \boldsymbol{\mu}(N_\mathbb{T}) &= \boldsymbol{\mu}\left((\mathbb{x}, \mathbb{T})^{-1}(\mathbf{R}^{2d} \setminus \text{spt } \gamma_\mathbb{T})\right) = \gamma_\mathbb{T}(\mathbf{R}^{2d} \setminus \text{spt } \gamma_\mathbb{T}) = 0 \end{aligned}$$

since $\gamma_\mathbb{T}$ is inner regular (being a finite Borel measure on a locally compact Hausdorff space with countable basis; see [37]). As $\text{spt } \gamma_\mathbb{T}$ is monotone, condition (4.6) follows.

Conversely, suppose $\mathbb{T} \in \mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu})$ satisfies (4.6). Let $\gamma_\mathbb{T} := (\mathbb{x}, \mathbb{T}) \# \boldsymbol{\mu}$. For every $(x_i, y_i) \in \text{spt } \gamma_\mathbb{T}$ with $i = 1..2$ and any $\varepsilon > 0$ we have

$$\begin{aligned} 0 &< \gamma_\mathbb{T}(B_\varepsilon(x_i) \times B_\varepsilon(y_i)) \\ &= \boldsymbol{\mu}\left(\left\{ (x, \xi) \in \mathbf{R}^{2d} : (x, \mathbb{T}(x, \xi)) \in B_\varepsilon(x_i) \times B_\varepsilon(y_i) \right\}\right), \end{aligned}$$

by definition of support. Therefore there exist $(\hat{x}_i, \hat{\xi}_i) \in \mathbf{R}^{2d}$ such that

$$(\hat{x}_i, \mathbb{T}(\hat{x}_i, \hat{\xi}_i)) \in B_\varepsilon(x_i) \times B_\varepsilon(y_i) \quad \text{for } i = 1..2.$$

We may assume that $(\hat{x}_i, \hat{\xi}_i) \notin N_\mathbb{T}$, where $N_\mathbb{T}$ is the null set in (4.6). Then

$$\langle y_1 - y_2, x_1 - x_2 \rangle \geq \langle \mathbb{T}(\hat{x}_1, \hat{\xi}_1) - \mathbb{T}(\hat{x}_2, \hat{\xi}_2), \hat{x}_1 - \hat{x}_2 \rangle - M\varepsilon,$$

with $M := 4 \max_{i=1..2} \{|x_i|, |y_i|\} + 2\varepsilon$. Since $(x_i, y_i) \in \text{spt } \gamma_\mathbb{T}$ and $\varepsilon > 0$ are arbitrary, we conclude that $\text{spt } \gamma_\mathbb{T}$ is monotone, and therefore $\mathbb{T} \in \mathcal{C}_\mu$.

The cone \mathcal{C}_μ is the set of all possible configurations, with reference configuration determined by $\boldsymbol{\mu} \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$. We will occasionally refer to \mathcal{C}_μ as the configuration manifold. By construction, configurations do not allow interpenetration of matter since the maps are monotone. They do admit, however, the concentration of mass if the transport map is not strictly monotone. The fluid element at location/velocity $(x, \xi) \in \mathbf{R}^d \times \mathbf{R}^d$ will never split because its final position is a *function* of (x, ξ) .

Definition 4.5 (Tangent Cone). Let $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ and $\boldsymbol{\mu} \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ be given. The tangent cone of \mathcal{C}_μ at the configuration $\mathbb{T} \in \mathcal{C}_\mu$ is defined as

$$\text{Tan}_\mathbb{T} \mathcal{C}_\mu := \overline{\left\{ \mathbb{V} \in \mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu}) : \text{there exists } \varepsilon > 0 \text{ with } \mathbb{T} + \varepsilon \mathbb{V} \in \mathcal{C}_\mu \right\}}^{\mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu})}.$$

This is the usual definition for tangent cones (also called support cones) to closed convex sets in Hilbert spaces; see §2 in [64]. Hence $\text{Tan}_\mathbb{T} \mathcal{C}_\mu$ is a closed convex cone with vertex at the origin (that is, the zero map) containing $\mathcal{C}_\mu - \mathbb{T}$. We also consider the dual (also called: polar) cone to $\text{Tan}_\mathbb{T} \mathcal{C}_\mu$, which is defined by

$$\begin{aligned} \text{Tan}_\mathbb{T}^* \mathcal{C}_\mu &:= \left\{ \mathbb{W} \in \mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu}) : \right. \\ &\quad \left. \int_{\mathbf{R}^{2d}} \langle \mathbb{W}(x, \xi), \mathbb{V}(x, \xi) \rangle \boldsymbol{\mu}(dx, d\xi) \leq 0 \quad \text{for all } \mathbb{V} \in \text{Tan}_\mathbb{T} \mathcal{C}_\mu \right\} \end{aligned}$$

and a closed convex cone with vertex at the origin as well. The cones $\text{Tan}_\mathbb{T} \mathcal{C}_\mu$ and $\text{Tan}_\mathbb{T}^* \mathcal{C}_\mu$ are duals of each other (see Lemma 2.1 in [64]); their only shared element

is the origin. For any $\mathbb{W} \in \mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu})$ there exists a unique metric projection onto the closed convex cone $\text{Tan}_{\mathbb{T}} \mathcal{C}_{\boldsymbol{\mu}}$, which we will denote by $\mathbb{P}_{\text{Tan}_{\mathbb{T}} \mathcal{C}_{\boldsymbol{\mu}}}(\mathbb{W})$. It can be characterized by the following property: for all $\mathbb{U} \in \mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu})$ we have

$$\mathbb{U} = \mathbb{P}_{\text{Tan}_{\mathbb{T}} \mathcal{C}_{\boldsymbol{\mu}}}(\mathbb{W}) \quad \text{if and only if}$$

$$\int_{\mathbf{R}^{2d}} \langle \mathbb{W}(x, \xi) - \mathbb{U}(x, \xi), \mathbb{U}(x, \xi) - \mathbb{V}(x, \xi) \rangle \boldsymbol{\mu}(dx, d\xi) \geq 0 \quad \text{for all } \mathbb{V} \in \text{Tan}_{\mathbb{T}} \mathcal{C}_{\boldsymbol{\mu}}.$$

Since $\text{Tan}_{\mathbb{T}} \mathcal{C}_{\boldsymbol{\mu}}$ is a cone, the latter condition is equivalent to

$$\begin{aligned} \int_{\mathbf{R}^{2d}} \langle \mathbb{W}(x, \xi) - \mathbb{U}(x, \xi), \mathbb{U}(x, \xi) \rangle \boldsymbol{\mu}(dx, d\xi) &= 0 \\ \int_{\mathbf{R}^{2d}} \langle \mathbb{W}(x, \xi) - \mathbb{U}(x, \xi), \mathbb{V}(x, \xi) \rangle \boldsymbol{\mu}(dx, d\xi) &\leq 0 \quad \text{for all } \mathbb{V} \in \text{Tan}_{\mathbb{T}} \mathcal{C}_{\boldsymbol{\mu}}. \end{aligned}$$

We will also need the following fact: for any $\mathbb{T} \in \mathcal{C}_{\boldsymbol{\mu}}$ we have

$$\|\mathbb{V}\|_{\mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu})}^{-1} \left\| \left(\mathbb{P}_{\mathcal{C}_{\boldsymbol{\mu}}}(\mathbb{T} + \mathbb{V}) - \mathbb{T} \right) - \mathbb{P}_{\text{Tan}_{\mathbb{T}} \mathcal{C}_{\boldsymbol{\mu}}}(\mathbb{V}) \right\|_{\mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu})} \longrightarrow 0 \quad (4.7)$$

as $\mathbb{V} \rightarrow 0$ over any locally compact cone of increments; see Lemma 4.6 of [64]. The metric projection $\mathbb{P}_{\text{Tan}_{\mathbb{T}} \mathcal{C}_{\boldsymbol{\mu}}}$ is therefore the differential of $\mathbb{P}_{\mathcal{C}_{\boldsymbol{\mu}}}$ at $\mathbb{T} \in \mathcal{C}_{\boldsymbol{\mu}}$.

Let us collect some additional properties of the tangent cone.

Proposition 4.6 (Tangent Cone). *With $\mathbb{T} \in \mathcal{C}_{\boldsymbol{\mu}}$ and $\boldsymbol{\gamma}_{\mathbb{T}} := (\mathbb{x}, \mathbb{T}) \# \boldsymbol{\mu}$, assume*

$$\langle y_1 - y_2, x_1 - x_2 \rangle \geq \alpha |x_1 - x_2|^2 \quad \text{for all } (x_i, y_i) \in \text{spt } \boldsymbol{\gamma}_{\mathbb{T}} \quad (4.8)$$

and $i = 1..2$, where $\alpha > 0$ is some constant. Then we have:

- (i.) Every $\mathcal{V} \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$ is contained in $\text{Tan}_{\mathbb{T}} \mathcal{C}_{\boldsymbol{\mu}}$.
- (ii.) For all $\mathbb{V} \in \text{Tan}_{\mathbb{T}} \mathcal{C}_{\boldsymbol{\mu}}$ we also have $-\mathbb{V} \in \text{Tan}_{\mathbb{T}} \mathcal{C}_{\boldsymbol{\mu}}$.

In particular, the tangent cone $\text{Tan}_{\mathbb{T}} \mathcal{C}_{\boldsymbol{\mu}}$ is a closed subspace of $\mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu})$.

Proof. We divide the proof into three steps.

Step 1. Note first that any finite Borel measure ν on a locally compact Hausdorff space Ω with continuous base is inner regular. Therefore the space of all continuous functions with compact support is dense in $\mathcal{L}^2(\Omega, \nu)$. We refer the reader to [37] for further details. If Ω is also a vector space, then the same statement is true for *smooth* functions with compact support. For any $\mathcal{V} \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$ there exists thus a sequence of smooth functions \mathcal{V}^m with compact support with $\mathcal{V}^m \rightarrow \mathcal{V}$ strongly in $\mathcal{L}^2(\mathbf{R}^d, \varrho)$ and therefore in $\mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu})$. We claim that $\mathbb{T} + \varepsilon \mathcal{V}^m \in \mathcal{C}_{\boldsymbol{\mu}}$ for $\varepsilon > 0$ sufficiently small. To prove this, let $N_{\mathbb{T}}$ be the null set of Remark 4.4. Then

$$\begin{aligned} &\langle (\mathbb{T}(x_1, \xi_1) + \varepsilon \mathcal{V}^m(x_1)) - (\mathbb{T}(x_2, \xi_2) + \varepsilon \mathcal{V}^m(x_2)), x_1 - x_2 \rangle \\ &= \langle \mathbb{T}(x_1, \xi_1) - \mathbb{T}(x_2, \xi_2), x_1 - x_2 \rangle - \varepsilon \langle \mathcal{V}^m(x_1) - \mathcal{V}^m(x_2), x_1 - x_2 \rangle \\ &\geq \left(\alpha - \varepsilon \|D\mathcal{V}^m\|_{\mathcal{L}^\infty(\mathbf{R}^d)} \right) |x_1 - x_2|^2 \geq 0 \end{aligned}$$

for all $(x_i, \xi_i) \in \mathbf{R}^{2d} \setminus N_{\mathbb{T}}$ with $i = 1..2$, for $\varepsilon > 0$ small. This proves part (i.).

Step 2. We now show that for every $\mathbb{S} \in \mathcal{C}_{\boldsymbol{\mu}}$ we also have $-\mathbb{S} \in \text{Tan}_{\mathbb{T}} \mathcal{C}_{\boldsymbol{\mu}}$. The argument is a modification of the proof of Proposition 4.28 of [41]. We first define the plan $\boldsymbol{\gamma} := (\mathbb{x}, \mathbb{S}) \# \boldsymbol{\mu} \in C_{\varrho}$. Since $\text{spt } \boldsymbol{\gamma}$ is a monotone subset of $\mathbf{R}^d \times \mathbf{R}^d$, there

exists a maximal monotone extension of it, which we denote by Γ . Let u be the corresponding maximal monotone set-valued map, defined as

$$u(x) := \{y \in \mathbf{R}^d : (x, y) \in \Gamma\} \quad \text{for all } x \in \mathbf{R}^d.$$

It is well-known that for every $x \in \mathbf{R}^d$ the image $u(x)$ is a closed and convex subset of \mathbf{R}^d ; see [1]. Consider the disintegration of the transport plan

$$\gamma(dx, dy) =: \gamma_x(dy) \varrho(dx).$$

Then we have that $\gamma_x = \mathbb{S}(x, \cdot) \# \mu_x$ for ϱ -a.e. $x \in \mathbf{R}^d$, with $\mu(dx, d\xi) = \mu_x(d\xi) \varrho(dx)$ the disintegration of μ . Let $\check{\gamma}_x := (-\mathbb{S}(x, \cdot)) \# \mu_x$ and $-\gamma = (\mathbb{S}, -\mathbb{S}) \# \mu$ so that

$$(-\gamma)(dx, dy) = \check{\gamma}_x(dy) \varrho(dx).$$

For ϱ -a.e. $x \in \mathbf{R}^d$ we denote by $A_x \subset u(x)$ the closed convex hull of $\text{spt } \gamma_x$. For such a x there are two possibilities: either γ_x is a Dirac measure and $A_x = \{\mathbb{B}(\gamma)(x)\}$ (recall Definition 2.5), or A_x (and therefore $u(x)$) contains $\mathbb{B}(\gamma)(x)$ as an interior point with respect to the relative topology. In the latter case, the subspace

$$L_x := \bigcup_{n \in \mathbf{N}} n(-\mathbb{B}(\gamma)(x) + u(x))$$

has the property that $\text{spt } \gamma_x \subset \mathbb{B}(\gamma)(x) + L_x$, and hence $\text{spt } \check{\gamma}_x \subset -\mathbb{B}(\gamma)(x) + L_x$. Let \mathbb{P}_x^n be the metric projection of \mathbf{R}^d onto the closed convex set

$$-(n+1)\mathbb{B}(\gamma)(x) + nu(x) \tag{4.9}$$

for ϱ -a.e. $x \in \mathbf{R}^d$ and $n \in \mathbf{N}$. Since projections are contractions, we have

$$\begin{aligned} |\mathbb{P}_x^n(y) - y| &\leq |\mathbb{P}_x^n(y) - \mathbb{P}_x^n(-\mathbb{B}(\gamma)(x))| + |(-\mathbb{B}(\gamma)(x)) - y| \\ &\leq 2|y - (-\mathbb{B}(\gamma)(x))| \end{aligned}$$

for all $y \in \mathbf{R}^d$. We used that $-\mathbb{B}(\gamma)(x)$ is contained in (4.9). We have

$$\begin{aligned} &\int_{\mathbf{R}^d} \left(\int_{\mathbf{R}^d} |y - (-\mathbb{B}(\gamma)(x))|^2 \check{\gamma}_x(dy) \right) \varrho(dx) \\ &\leq 4 \int_{\mathbf{R}^d} \left(\int_{\mathbf{R}^d} |y|^2 \check{\gamma}_x(dy) \right) \varrho(dx) = 4 \int_{\mathbf{R}^{2d}} |\mathbb{S}(x, \xi)|^2 \mu(dx, d\xi), \end{aligned}$$

by definition of $\mathbb{B}(\gamma)(x)$ and Jensen's inequality. We now define the maps

$$\mathbb{S}^n(x, \xi) := \mathbb{P}_x^n(-\mathbb{S}(x, \xi)) \quad \text{for } \mu\text{-a.e. } (x, \xi) \in \mathbf{R}^{2d}.$$

Using dominated convergence, we get for $n \rightarrow \infty$ that

$$\|\mathbb{S}^n - (-\mathbb{S})\|_{\mathcal{L}^2(\mathbf{R}^{2d}, \mu)}^2 = \int_{\mathbf{R}^d} \left(\int_{\mathbf{R}^d} |\mathbb{P}_x^n(y) - y|^2 \check{\gamma}_x(dy) \right) \varrho(dx) \longrightarrow 0, \tag{4.10}$$

because $\mathbb{P}_x^n(y) \longrightarrow y$ for all $y \in -\mathbb{B}(\gamma)(x) + L_x$ and for ϱ -a.e. $x \in \mathbf{R}^d$ such that γ_x is not a Dirac measure. We used again that $\mathbb{P}_x^n(-\mathbb{B}(\gamma)(x)) = -\mathbb{B}(\gamma)(x)$.

As discussed in Step 1, there exists a sequence of smooth, compactly supported functions \mathcal{Y}^m such that $\mathcal{Y}^m \longrightarrow \mathbb{B}(\gamma)$ in $\mathcal{L}^2(\mathbf{R}^d, \varrho)$. We now define

$$\mathbb{S}^{n,m}(x, \xi) := \mathbb{S}^n(x, \xi) + (n+1)(\mathbb{B}(\gamma)(x) - \mathcal{Y}^m(x)). \tag{4.11}$$

for μ -a.e. $(x, \xi) \in \mathbf{R}^{2d}$. We get for $m \rightarrow \infty$ (with n fixed) that

$$\|\mathbb{S}^{n,m} - \mathbb{S}^n\|_{\mathcal{L}^2(\mathbf{R}^{2d}, \mu)}^2 = (n+1)^2 \int_{\mathbf{R}^d} |\mathbb{B}(\gamma)(x) - \mathcal{Y}^m(x)|^2 \varrho(dx) \longrightarrow 0. \tag{4.12}$$

Combining (4.10) and (4.12), we find $\|\mathbb{S}^{n,m} - (-\mathbb{S})\|_{\mathcal{L}^2(\mathbf{R}^{2d}, \mu)} \rightarrow 0$. We claim that $\mathbb{T} + \varepsilon \mathbb{S}^{n,m} \in \mathcal{C}_\mu$ for $\varepsilon > 0$ small. To prove this, we observe first that

$$\mathbb{S}^{n,m}(x, \xi) \subset -(n+1)\mathcal{Y}^m(x) + nu(x) \quad \text{for } \mu\text{-a.e. } (x, \xi) \in \mathbf{R}^{2d},$$

by definition of \mathbb{P}_x^n and (4.11). With $N_{\mathbb{T}}$ the null set of Remark 4.4, we have

$$\begin{aligned} & \langle (\mathbb{T}(x_1, \xi_1) - \varepsilon(n+1)\mathcal{Y}^m(x_1)) - (\mathbb{T}(x_2, \xi_2) - \varepsilon(n+1)\mathcal{Y}^m(x_2)), x_1 - x_2 \rangle \\ &= \langle \mathbb{T}(x_1, \xi_1) - \mathbb{T}(x_1, \xi_1), x_1 - x_2 \rangle - \varepsilon(n+1) \langle \mathcal{Y}^m(x_1) - \mathcal{Y}^m(x_2), x_1 - x_2 \rangle \\ &\geq \left(\alpha - \varepsilon(n+1) \|D\mathcal{Y}^m\|_{\mathcal{L}^\infty(\mathbf{R}^d)} \right) |x_1 - x_2|^2 \geq 0 \end{aligned}$$

for all $(x_i, \xi_i) \in \mathbf{R}^{2d} \setminus N_{\mathbb{T}}$ with $i = 1..2$, for $\varepsilon > 0$ small. Since u is monotone, the support of $(\mathbb{x}, \mathbb{T} + \varepsilon \mathbb{S}^{n,m}) \# \mu$ is contained in a monotone subset of $\mathbf{R}^d \times \mathbf{R}^d$.

Step 3. We prove that if $\mathbb{V} \in \text{Tan}_{\mathbb{T}} \mathcal{C}_\mu$ then also $-\mathbb{V} \in \text{Tan}_{\mathbb{T}} \mathcal{C}_\mu$. There exists a sequence of $\mathbb{V}^n \in \mathcal{L}^2(\mathbf{R}^{2d}, \mu)$ with $\|\mathbb{V}^n - \mathbb{V}\|_{\mathcal{L}^2(\mathbf{R}^{2d}, \mu)} \rightarrow 0$ as $n \rightarrow \infty$, and such that $\mathbb{T} + \varepsilon^n \mathbb{V}^n \in \mathcal{C}_\mu$ for $\varepsilon^n > 0$ small. We have the following identity:

$$-\mathbb{V}^n = -\frac{1}{\varepsilon^n}(\mathbb{T} + \varepsilon^n \mathbb{V}^n) + \frac{1}{\varepsilon^n} \mathbb{T}.$$

The first term on the right-hand side is in $\text{Tan}_{\mathbb{T}} \mathcal{C}_\mu$ because of Step 2; the second one is in $\mathcal{C}_\mu \subset \text{Tan}_{\mathbb{T}} \mathcal{C}_\mu$. Since the tangent cone is a closed convex cone, we conclude that $-\mathbb{V}^n \in \text{Tan}_{\mathbb{T}} \mathcal{C}_\mu$. Then we use that $\|(-\mathbb{V}^n) - (-\mathbb{V})\|_{\mathcal{L}^2(\mathbf{R}^{2d}, \mu)} \rightarrow 0$. \square

Remark 4.7. We emphasize that, unlike the tangent cone built from optimal transport maps/plans, which basically consists of *gradient* vector fields (see [4, 41]), the tangent cone derived from monotone maps contains all of $\mathcal{L}^2(\mathbf{R}^d, \varrho)$, provided \mathbb{T} is strictly monotone in the sense of (4.8). This condition is satisfied when $\mathbb{T} = \mathbb{x}$, for example. Otherwise, the tangent cone may again be a proper subset: If $d = 1$ and $\mathcal{T} \in \mathcal{C}_\mu$ depends only on the spatial variable $x \in \mathbf{R}$ (which implies $\mathcal{T} \in \mathcal{L}^2(\mathbf{R}, \varrho)$), then a map $\mathcal{V} \in \mathcal{L}^2(\mathbf{R}, \varrho)$ can only be in $\text{Tan}_{\mathcal{T}} \mathcal{C}_\mu$ when \mathcal{V} is non-decreasing on each open interval on which \mathcal{T} is constant; see Lemma 3.6 in [16].

4.2. Minimization Problem. We now introduce the main minimization problem for (1.3). For sticky particle solutions, all fluid elements meeting at the same location stick together and form a larger compound. Therefore the measures $\mu \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$ describing the state of the fluid must be monokinetic ϱ -a.e. Mass and momentum are conserved. Since transport maps $\mathbb{T} \in \mathcal{C}_\mu$ are not required to be strictly monotone (hence injective), it may happen that fluid elements with distinct velocities are transported to the same location. We will then use the barycentric projection to define a single velocity. This implements the sticky particle condition.

Definition 4.8. For any $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ and $\mu \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$, and any $\mathbb{T} \in \mathcal{C}_\mu$ let

$$\mathcal{H}_\mu(\mathbb{T}) := \left\{ \mathbf{u} \circ \mathbb{T} : \mathbf{u} \in \mathcal{L}^2(\mathbf{R}^d, \varrho_{\mathbb{T}}) \right\}, \quad \varrho_{\mathbb{T}} := \mathbb{T} \# \mu.$$

One can check that $\mathcal{H}_\mu(\mathbb{T})$ is a closed subspace of $\mathcal{L}^2(\mathbf{R}^{2d}, \mu)$ because

$$\|\mathbf{u} \circ \mathbb{T}\|_{\mathcal{L}^2(\mathbf{R}^{2d}, \mu)} = \|\mathbf{u}\|_{\mathcal{L}^2(\mathbf{R}^d, \varrho_{\mathbb{T}})}$$

for all $\mathbf{u} \circ \mathbb{T} \in \mathcal{H}_\mu(\mathbb{T})$; see Section 5.2 in [4]. Consequently, there exists an orthogonal projection onto this subspace, which we will denote by $\mathbb{P}_{\mathcal{H}_\mu(\mathbb{T})}$.

Definition 4.9 (Energy Minimization). Let $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ and $\boldsymbol{\mu} \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$, and $\tau > 0$ be given. Then we consider the following three-step scheme:

- (1) Compute the metric projection $\mathbb{T}_\tau := \mathbb{P}_{\mathcal{C}_\mu}(\mathbb{x} + \tau \mathbb{v})$ and define

$$\mathbb{W}_\tau(x, \xi) := V_\tau(x, \xi, \mathbb{T}_\tau(x, \xi)) \quad \text{for } \boldsymbol{\mu}\text{-a.e. } (x, \xi) \in \mathbf{R}^{2d}, \quad (4.13)$$

see Remark 4.3 and (3.6) for the definition of V_τ .

- (2) Compute the orthogonal projection $\mathbb{U}_\tau := \mathbb{P}_{\mathcal{H}_\mu(\mathbb{T}_\tau)}(\mathbb{W}_\tau)$.
- (3) Define the updated fluid state

$$\varrho_\tau := \mathbb{T}_\tau \# \boldsymbol{\mu}, \quad \boldsymbol{\mu}_\tau := (\mathbb{T}_\tau, \mathbb{U}_\tau) \# \boldsymbol{\mu}.$$

We emphasize the fact that the monotone transport map \mathbb{T}_τ and the intermediate velocity \mathbb{W}_τ are connected by (4.13). Defining also the transport velocity

$$\mathbb{V}_\tau(x, \xi) := \frac{\mathbb{T}_\tau(x, \xi) - x}{\tau} \quad \text{for } \boldsymbol{\mu}\text{-a.e. } (x, \xi) \in \mathbf{R}^{2d},$$

we have the following identities, which will be used frequently:

$$(x + \tau \xi) - \mathbb{T}_\tau(x, \xi) = \tau(\xi - \mathbb{V}_\tau(x, \xi)) = \frac{2\tau}{3}(\xi - \mathbb{W}_\tau(x, \xi)).$$

In particular, the transport velocity can be written as a convex combination

$$\mathbb{V}_\tau(x, \xi) = \frac{2}{3}\mathbb{W}_\tau(x, \xi) + \frac{1}{3}\xi \iff \mathbb{W}_\tau(x, \xi) = \frac{3}{2}\mathbb{V}_\tau(x, \xi) - \frac{1}{2}\xi \quad (4.14)$$

for $\boldsymbol{\mu}$ -a.e. $(x, \xi) \in \mathbf{R}^{2d}$. Notice that if $\mathbf{u}_\tau \in \mathcal{L}^2(\mathbf{R}^d, \varrho_\tau)$ is determined by the identity $\mathbb{U}_\tau =: \mathbf{u}_\tau \circ \mathbb{T}_\tau$, then we can write $\boldsymbol{\mu}_\tau = (\text{id}, \mathbf{u}_\tau) \# \varrho_\tau$. We also observe that \mathbf{u}_τ is just the barycentric projection of $\boldsymbol{\mu}_* := (\mathbb{T}_\tau, \mathbb{W}_\tau) \# \boldsymbol{\mu}$: We have

$$\int_{\mathbf{R}^{2d}} |\mathbb{W}_\tau(x, \xi) - \mathbf{u}(\mathbb{T}_\tau(x, \xi))|^2 \boldsymbol{\mu}(dx, d\xi) = \int_{\mathbf{R}^{2d}} |\zeta - \mathbf{u}(z)|^2 \boldsymbol{\mu}_*(dz, d\zeta)$$

for all $\mathbf{u} \in \mathcal{L}^2(\mathbf{R}^d, \varrho_\tau)$, and the barycentric projection $\mathbb{B}(\boldsymbol{\mu}_*)$ is the unique element in $\mathcal{L}^2(\mathbf{R}^d, \varrho_\tau)$ closest to $\boldsymbol{\mu}_*$ with respect to $\mathbb{W}_{\varrho_\tau}$ (recall (2.2)). The discussion in the previous sections shows that $\boldsymbol{\mu}_\tau$ is well-defined for every choice of $(\varrho, \boldsymbol{\mu}, \tau)$.

From Proposition 4.6, we get that $\mathcal{L}^2(\mathbf{R}^d, \varrho_\tau) \subset \text{Tan}_{\mathbb{x}} \mathcal{C}_{\boldsymbol{\mu}_*}$. Therefore Step (2) of Definition 4.9 may be interpreted as the projection of the updated state $\boldsymbol{\mu}_*$ onto (a subspace of) the tangent cone at the new configuration. A similar combination of transporting the vector field, then projecting it onto the tangent cone was used in [3] to construct the parallel transport along curves in $\mathcal{P}_2(\mathbf{R}^D)$; see also [11]. Notice that when ϱ_τ is absolutely continuous with respect to the Lebesgue measure so that there is no concentration (no sticking together of fluid elements), then the tangent cone at ϱ_τ consists only of monokinetic states, as follows from Remark 3.14.

Remark 4.10. Since the metric projection $\mathbb{P}_{\mathcal{C}_\mu}$ is a contraction (see [64]), we have that $\mathbb{T}_\tau \rightarrow \mathbb{x}$ strongly in $\mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu})$ as $\tau \rightarrow 0$. On the other hand, because of the second identity in (4.14) we obtain $\mathbb{W}_\tau \rightarrow \mathbb{v}$ strongly in $\mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu})$ only if $\mathbb{v} \in \text{Tan}_{\mathbb{x}} \mathcal{C}_\mu$; see (4.7). This is the case e.g. if $\boldsymbol{\mu} = (\text{id}, \mathbf{u}) \# \varrho$ for some $\mathbf{u} \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$, because of Proposition 4.6 (i.). The convergence of \mathbb{U}_τ will be studied elsewhere.

Remark 4.11. In [14] the authors prove the non-existence of sticky particle solutions to (1.3) for well-designed initial data. In their construction the number of collisions grows unboundedly the closer one gets to the initial time, and so the dynamics has arbitrarily small time scales. Using our discretization, we can construct a sequence of approximate solutions to (1.3) starting from the initial data in [14]. We will show

below that this approximation converges to a measure-valued solution of (1.3). The timestep $\tau > 0$ in our discretization introduces a minimal time scale below which the dynamics is not completely resolved but is “smeared out.” It would be interesting to know to which solution our discretization converges in the limit $\tau \rightarrow 0$.

Remark 4.12. The constant map $\mathbb{S}(x, \xi) = \pm b$ for all $(x, \xi) \in \mathbf{R}^{2d}$, where $b \in \mathbf{R}^d$ is some vector, is an element of \mathcal{C}_μ . Using this function in (4.5), we get

$$\begin{aligned} \int_{\mathbf{R}^{2d}} \langle b, \zeta \rangle \mu_\tau(dz, d\zeta) &= \int_{\mathbf{R}^{2d}} \langle b, \mathbb{U}_\tau(x, \xi) \rangle \mu(dx, d\xi) \\ &= \int_{\mathbf{R}^{2d}} \langle b, \mathbb{W}_\tau(x, \xi) \rangle \mu(dx, d\xi) = \int_{\mathbf{R}^{2d}} \langle b, \xi \rangle \mu(dx, d\xi); \end{aligned}$$

see (4.14). Recall that \mathbb{U}_τ is the orthogonal projection of \mathbb{W}_τ onto $\mathcal{H}_\mu(\mathbb{T})$, which contains \mathbb{S} . We conclude that the minimization preserves the total momentum.

Remark 4.13. Let us assume that $d = 1$ and $\mu = (\text{id}, \mathbf{u}) \# \varrho$ for some $\mathbf{u} \in \mathcal{L}^2(\mathbf{R}, \varrho)$. We use the notation of Definition 4.9 and the subsequent discussion. Then

$$\begin{aligned} \int_{\mathbf{R}^2} \langle \xi - \mathbf{u}_\tau(\mathbb{T}_\tau(x, \xi)), \mathbf{v}(\mathbb{T}_\tau(x, \xi)) \rangle \mu(dx, d\xi) \\ = \int_{\mathbf{R}^2} \langle \xi - \mathbb{W}_\tau(x, \xi), \mathbf{v}(\mathbb{T}_\tau(x, \xi)) \rangle \mu(dx, d\xi) \end{aligned}$$

for all $\mathbf{v} \in \mathcal{D}(\mathbf{R})$. Using the identity (4.14) for \mathbb{W}_τ , we can write

$$\begin{aligned} \int_{\mathbf{R}^2} \langle \xi - \mathbb{W}_\tau(x, \xi), \mathbf{v}(\mathbb{T}_\tau(x, \xi)) \rangle \mu(dx, d\xi) \\ = \frac{3}{2\tau} \int_{\mathbf{R}^2} \langle (x + \tau\xi) - \mathbb{T}_\tau(x, \xi), (\mathbf{v} + \alpha \text{id}) \circ \mathbb{T}_\tau(x, \xi) \rangle \mu(dx, d\xi) \\ - \frac{3\alpha}{2\tau} \int_{\mathbf{R}^2} \langle (x + \tau\xi) - \mathbb{T}_\tau(x, \xi), \mathbb{T}_\tau(x, \xi) \rangle \mu(dx, d\xi) \end{aligned}$$

for any $\alpha > 0$. The last integral vanishes because of (4.4) in Remark 4.3. Define now $\mathcal{T}_\tau(x) := \mathbb{T}_\tau(x, \mathbf{u}(x))$ for ϱ -a.e. $x \in \mathbf{R}$. For α large enough the map $\mathbf{v} + \alpha \text{id}$ is strictly increasing. Then one can check that $(\text{id}, (\mathbf{v} + \alpha \text{id}) \circ \mathcal{T}_\tau) \# \varrho \in C_\varrho$. Here we used the assumption that $d = 1$ (since compositions of monotone maps are again monotone in one space dimension). Using inequality (4.5), we have

$$\int_{\mathbf{R}^2} \langle \xi - \mathbb{W}_\tau(x, \xi), \mathbf{v}(\mathbb{T}_\tau(x, \xi)) \rangle \mu(dx, d\xi) \leq 0 \quad \text{for all } \mathbf{v} \in \mathcal{D}(\mathbf{R}),$$

which in particular implies equality. Since $\mathcal{D}(\mathbf{R})$ is dense in $\mathcal{L}^2(\mathbf{R}, \varrho_\tau)$ we get

$$\int_{\mathbf{R}} \langle \mathbf{u}(x) - \mathbf{u}_\tau(\mathcal{T}_\tau(x)), \mathbf{v}(\mathcal{T}_\tau(x)) \rangle \varrho(dx) = 0 \quad \text{for all } \mathbf{v} \in \mathcal{L}^2(\mathbf{R}, \varrho_\tau).$$

Recall $\mu = (\text{id}, \mathbf{u}) \# \varrho$. Hence \mathbf{u}_τ is uniquely determined as the orthogonal projection of \mathbf{u} onto the closed subspace of $\mathcal{L}^2(\mathbf{R}, \varrho)$ consisting of functions of the form $\mathbf{v} \circ \mathcal{T}_\tau$ with $\mathbf{v} \in \mathcal{L}^2(\mathbf{R}, \varrho_\tau)$. But this property, in combination with the definition of \mathcal{T}_τ as the metric projection of $\text{id} + \tau \mathbf{u}$ onto monotone maps, characterizes the solutions of (1.3), for a.e. $\tau > 0$. So our discretization already generates the *exact* solution, not just an approximation, for $d = 1$. We refer the reader to [16] for details.

4.3. Polar Cone. In this section, we will give a representation of the elements in the polar cone of \mathcal{C}_μ . As we will see later, such elements appear as stress tensors. Let $\mathcal{C}_*(\mathbf{R}^d; \mathbf{R}^D)$ be the space of all continuous functions $U: \mathbf{R}^d \rightarrow \mathbf{R}^D$ with the property that $\lim_{|x| \rightarrow \infty} U(x) \in \mathbf{R}^D$ exists. Note that we can write

$$\mathcal{C}_*(\mathbf{R}^d; \mathbf{R}^D) = \mathbf{R}^D + \mathcal{C}_0(\mathbf{R}^d; \mathbf{R}^D),$$

where $\mathcal{C}_0(\mathbf{R}^d; \mathbf{R}^D)$ is the closure of the space of all compactly supported continuous \mathbf{R}^D -valued maps w.r.t. the sup-norm. We can identify $\mathcal{C}_*(\mathbf{R}^d; \mathbf{R}^D)$ with the space $\mathcal{C}(\gamma\mathbf{R}^d; \mathbf{R}^D)$ of continuous functions on the one-point compactification $\gamma\mathbf{R}^d$ of \mathbf{R}^d . We adjoin to \mathbf{R}^d an additional point called ∞ and define a distance (see [52])

$$d(x, y) := \begin{cases} \min\{|x - y|, h(x) + h(y)\} & \text{if } x, y \in \mathbf{R}^d, \\ h(x) & \text{if } x \in \mathbf{R}^d \text{ and } y = \infty, \\ 0 & \text{if } x, y = \infty, \end{cases} \quad (4.15)$$

where $h(x) := 1/(1+|x|)$ for all $x \in \mathbf{R}^d$. Then $|x| \rightarrow \infty$ is equivalent to $d(x, \infty) \rightarrow 0$. To any $U \in \mathcal{C}_*(\mathbf{R}^d; \mathbf{R}^D)$ we associate $\gamma U \in \mathcal{C}(\gamma\mathbf{R}^d; \mathbf{R}^D)$ defined as

$$(\gamma U)(x) := \begin{cases} U(x) & \text{if } x \in \mathbf{R}^d, \\ \lim_{|x| \rightarrow \infty} U(x) & \text{if } x = \infty. \end{cases}$$

Conversely, the restriction of any function in $\mathcal{C}(\gamma\mathbf{R}^d; \mathbf{R}^D)$ to \mathbf{R}^d induces a function in $\mathcal{C}_*(\mathbf{R}^d; \mathbf{R}^D)$. We will hence not distinguish between the two spaces. Similarly, we define $\mathcal{C}_*(\mathbf{R}^d; \mathbf{R}^{l \times l})$, $\mathcal{C}_*(\mathbf{R}^d; \mathcal{S}^l)$, and $\mathcal{C}_*(\mathbf{R}^d; \mathcal{S}_+^l)$, with \mathcal{S}^l the space of symmetric $(l \times l)$ -matrices and \mathcal{S}_+^l its subspace of matrices that are positive semidefinite.

For any map $u \in \mathcal{C}^1(\mathbf{R}^d; \mathbf{R}^d)$ we denote by

$$\epsilon(u(x)) := \nabla u(x)^{\text{sym}} \quad \text{for all } x \in \mathbf{R}^d \quad (4.16)$$

its deformation tensor, which is an element of $\mathcal{C}(\mathbf{R}^d; \mathcal{S}^d)$. Let

$$\begin{aligned} \mathcal{C}_*(\mathbf{R}^d; \mathbf{R}^d) &:= \{u \in \mathcal{C}^1(\mathbf{R}^d; \mathbf{R}^d) : \nabla u \in \mathcal{C}_*(\mathbf{R}^d; \mathbf{R}^{d \times d})\}, \\ \text{MON}(\mathbf{R}^d) &:= \{u \in \mathcal{C}_*(\mathbf{R}^d; \mathbf{R}^d) : u \text{ is monotone}\}, \end{aligned}$$

so that $\epsilon(u) \in \mathcal{C}_*(\mathbf{R}^d; \mathcal{S}_+^d)$ if $u \in \text{MON}(\mathbf{R}^d)$. The cone $\text{MON}(\mathbf{R}^d)$ contains all linear maps $u(x) := Ax$ for $x \in \mathbf{R}^d$, with $A \in \mathbf{R}_+^{d \times d}$. Recall that the scalar product of two matrices $A, B \in \mathbf{R}^{d \times d}$ is defined as $\langle\langle A, B \rangle\rangle := \text{tr}(A^T B)$. Moreover, we have

$$|\langle\langle A, B \rangle\rangle| \leq \max_{i=1 \dots d} |\lambda_i| \left(\sum_{j=1}^d |\mu_j| \right)$$

whenever A, B are symmetric, with λ_i and μ_j the eigenvalues of A and B , respectively. The sum over $|\mu_j|$ equals the trace of the matrix if A is positive semidefinite, thus all eigenvalues are nonnegative. We will use the following result from [17]:

Theorem 4.14 (Stress Tensor). *Assume that there exists a measure $\mathbf{F} \in \mathcal{M}(\mathbf{R}^d; \mathbf{R}^d)$ with finite first moment and a matrix-valued field $\mathbf{H} \in \mathcal{M}(\mathbf{R}^d; \mathcal{S}_+^d)$ with*

$$G(u) := - \int_{\mathbf{R}^d} \langle u(x), \mathbf{F}(dx) \rangle - \int_{\mathbf{R}^d} \langle\langle \epsilon(u(x)), \mathbf{H}(dx) \rangle\rangle \geq 0 \quad (4.17)$$

for all $u \in \text{MON}(\mathbf{R}^d)$. There exists a measure $\mathbf{M} \in \mathcal{M}(\gamma\mathbf{R}^d; \mathcal{S}_+^d)$ such that

$$\begin{aligned} G(u) &= \int_{\gamma\mathbf{R}^d} \langle \epsilon(u(x)), \mathbf{M}(dx) \rangle \quad \text{for all } u \in \mathcal{C}_*^1(\mathbf{R}^d; \mathbf{R}^d), \\ \int_{\gamma\mathbf{R}^d} \text{tr}(\mathbf{M}(dx)) &= - \int_{\mathbf{R}^d} \langle x, \mathbf{F}(dx) \rangle - \int_{\mathbf{R}^d} \text{tr}(\mathbf{H}(dx)). \end{aligned} \quad (4.18)$$

Note that the integral in (4.17) is finite for any choice of $u \in \mathcal{C}_*^1(\mathbf{R}^d; \mathbf{R}^d)$ since the first moment of \mathbf{F} is finite, by assumption. Recall that the trace of a symmetric matrix equals the sum of its eigenvalues, which in the case of a positive semidefinite matrix are all nonnegative. Therefore (4.18) controls the size of \mathbf{M} .

Remark 4.15. The stress tensor \mathbf{M} does not actually assign any mass to the remainder $\gamma\mathbf{R}^d \setminus \mathbf{R}^d$, so Theorem 4.14 remains true if the compactification $\gamma\mathbf{R}^d$ is replaced by \mathbf{R}^d . In fact, recall that \mathbf{R}^d (being a separable metric space) is a Radon space, so that any finite Borel measure is inner regular. Consider a nonnegative, radially symmetric test function $\varphi \in \mathcal{D}(\mathbf{R}^d)$ with $\int_{\mathbf{R}^d} \varphi(x) dx = 1$ and define

$$u_R := \nabla(\phi_R \star \varphi), \quad \text{with} \quad \phi_R(x) := \frac{1}{2} \max\{|x|^2 - R^2, 0\}$$

for $x \in \mathbf{R}^d$ and $R > 0$. The map $\phi_R \star \varphi$ is convex and smooth (since the convolution preserves convexity), hence u_R is monotone and smooth. Notice that $u_R(x) = 0$ for all $|x| \leq R - c$, with c the (finite) diameter of $\text{spt } \varphi$. Moreover, we have

$$\int_{\mathbf{R}^d} \varphi(x - y) |y|^2 dy = |x|^2 + \left(\int_{\mathbf{R}^d} |z|^2 \varphi(z) dz \right)$$

for all $x \in \mathbf{R}^d$, which implies that $u_R(x) = x$ and $Du_R(x) = \mathbb{1}$ for $|x| \geq R + c$. In particular, we observe that $u_R \in \text{MON}(\mathbf{R}^d)$ for all $R > 0$. Then

$$\int_{|x| \geq R+c} \text{tr}(\mathbf{M}(dx)) \leq C \left(\int_{|x| \geq R-c} |x| |\mathbf{F}(dx)| + \int_{|x| \geq R-c} \text{tr}(\mathbf{H}(dx)) \right), \quad (4.19)$$

with C some finite constant depending on φ . The right-hand side of (4.19) converges to zero as $R \rightarrow \infty$ since both measures $|\mathbf{F}|$ and $\text{tr}(\mathbf{H})$ are inner regular and the first moment of \mathbf{F} is finite. We conclude that $\text{tr}(\mathbf{M})(\gamma\mathbf{R}^d \setminus \mathbf{R}^d) = 0$.

For $\mu \in \mathcal{P}_\rho(\mathbf{R}^{2d})$ and $\tau > 0$ let \mathbb{T}_τ be given by Definition 4.9. Then

$$-\frac{3}{2\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathbb{T}_\tau(x, \xi), \mathbb{S}(x, \xi) \rangle \mu(dx, d\xi) \geq 0 \quad \text{for all } \mathbb{S} \in \mathcal{C}_\mu,$$

which is (4.5). In particular, this inequality is true for $\mathbb{S} = u$ with $u \in \text{MON}(\mathbf{R}^d)$. Functions in $\text{MON}(\mathbf{R}^d)$ have at most linear growth and are therefore in $\mathcal{L}^2(\mathbf{R}^d, \rho)$. Applying Theorem 4.14 (with $\mathbf{H} \equiv 0$), we obtain $\mathbf{M}_\tau \in \mathcal{M}(\mathbf{R}^d; \mathcal{S}_+^d)$ with

$$\int_{\mathbf{R}^d} \langle \epsilon(u(x)), \mathbf{M}_\tau(dx) \rangle = -\frac{3}{2\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathbb{T}_\tau(x, \xi), u(x) \rangle \mu(dx, d\xi), \quad (4.20)$$

$$\int_{\mathbf{R}^d} \text{tr}(\mathbf{M}_\tau(dx)) = -\frac{3}{2\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathbb{T}_\tau(x, \xi), x \rangle \mu(dx, d\xi); \quad (4.21)$$

see Remark 4.15. In (4.20) we can replace $\epsilon(u(x))$ by $\nabla u(x)$ since \mathbf{M}_τ takes values in the symmetric matrices. The representation in Theorem 4.14 generalizes a similar

characterization of the polar cone of monotone maps obtained in [54] (in one space dimension). Using the identity (4.14), we obtain the following identities:

$$\begin{aligned} \int_{\mathbf{R}^d} \langle \epsilon(u(x)), \mathbf{M}_\tau(dx) \rangle &= -\frac{3}{2\tau} \int_{\mathbf{R}^{2d}} \langle \xi - \mathbb{V}_\tau(x, \xi), u(x) \rangle \mu(dx, d\xi) \\ &= -\frac{1}{\tau} \int_{\mathbf{R}^{2d}} \langle \xi - \mathbb{W}_\tau(x, \xi), u(x) \rangle \mu(dx, d\xi), \end{aligned} \quad (4.22)$$

with transport velocity $\mathbb{V}_\tau(x, \xi) := (\mathbb{T}_\tau(x, \xi) - x)/\tau$ for μ -a.e. $(x, \xi) \in \mathbf{R}^{2d}$, and

$$\begin{aligned} \int_{\mathbf{R}^d} \text{tr}(\mathbf{M}_\tau(dx)) &= -\frac{3}{2\tau} \int_{\mathbf{R}^{2d}} \langle \xi - \mathbb{V}_\tau(x, \xi), x \rangle \mu(dx, d\xi) \\ &= -\frac{1}{\tau} \int_{\mathbf{R}^{2d}} \langle \xi - \mathbb{W}_\tau(x, \xi), x \rangle \mu(dx, d\xi). \end{aligned}$$

Remark 4.16. In order to explore the significance of \mathbf{M}_τ , we consider

$$\mu(dx, d\xi) = \frac{1}{4} \delta_0(d\xi) \mathcal{L}^1|_{(-1,1)}(dx) + \frac{1}{2} \delta_1(d\xi) \delta_0(dx).$$

For any $\tau > 0$ the support of the transport plan $(\mathbb{x}, \mathbb{x} + \tau \mathbb{v}) \# \mu$ is not monotone. Then $\gamma_\tau := (\mathbb{x}, \mathbb{T}_\tau) \# \mu$, with \mathbb{T}_τ given by Definition 4.9, can be computed as

$$\begin{aligned} \gamma_\tau(dx, dy) &= \frac{1}{4} \left(\delta_{\beta(\tau)\tau}(dy) \mathcal{L}^1|_{[0, \beta(\tau)\tau]}(dx) + \delta_x(dy) \mathcal{L}^1|_{(-1,1) \setminus [0, \beta(\tau)\tau]}(dx) \right) \\ &\quad + \frac{1}{2} \delta_{\beta(\tau)\tau}(dy) \delta_0(dx), \end{aligned}$$

where $\beta(\tau) \in [0, 1]$ is the minimizer of the following function:

$$\varphi_\tau(\beta) := \frac{1}{2} |1 - \beta|^2 \tau^2 + \frac{1}{4} \int_0^{\beta\tau} |\beta\tau - x|^2 dx,$$

which represents the $\mathcal{L}^2(\mathbf{R}^{2d}, \mu)$ -distance of $\mathbb{x} + \tau \mathbb{v}$ to some map in \mathcal{C}_μ parameterized by β . One can check that $\beta(\tau) := \frac{2}{\tau}(\sqrt{1 + \tau} - 1)$ and $\beta(\tau) \rightarrow 1$ as $\tau \rightarrow 0$. The induced velocity distribution $\mu_\tau := (\mathbb{x}, (\mathbb{y} - \mathbb{x})/\tau) \# \gamma_\tau$ equals

$$\begin{aligned} \mu_\tau(dx, d\xi) &= \frac{1}{4} \left(\delta_{\beta(\tau) - x/\tau}(d\xi) \mathcal{L}^1|_{[0, \beta(\tau)\tau]}(dx) + \delta_0(d\xi) \mathcal{L}^1|_{(-1,1) \setminus [0, \beta(\tau)\tau]}(dx) \right) \\ &\quad + \frac{1}{2} \delta_{\beta(\tau)}(d\xi) \delta_0(dx). \end{aligned}$$

The first ξ -moments of μ and μ_τ determine the corresponding momenta:

$$\begin{aligned} \mathbf{m}(dx) &:= \frac{1}{2} \delta_0(dx), \\ \mathbf{m}_\tau(dx) &:= \frac{1}{4} \left(\beta(\tau) - \frac{x}{\tau} \right) \mathcal{L}^1|_{[0, \beta(\tau)\tau]}(dx) + \frac{1}{2} \beta(\tau) \delta_0(dx). \end{aligned}$$

Therefore the change in momenta (which represents an acceleration) has two parts: The velocity of the fluid element with mass $1/2$ located at $x = 0$ decreases, so the momentum is getting smaller. This momentum is *transferred* to fluid elements in the interval $[0, \beta(\tau)\tau]$, which pick up speed. The transfer is described by the derivative of the nonnegative measure from Theorem 4.14. Let $\mathbf{M}_\tau := M_\tau \mathcal{L}^1$ with

$$M_\tau(x) := \begin{cases} \frac{1}{2}(1 - \beta(\tau)) - \frac{1}{4} \left(\beta(\tau)x - \frac{x^2}{2\tau} \right) & \text{if } x \in [0, \beta(\tau)\tau], \\ 0 & \text{otherwise.} \end{cases}$$

Since $\frac{1}{2}(1 - \beta(\tau)) = \frac{1}{8}\beta(\tau)^2\tau$ the measure \mathbf{M}_τ is nonnegative, supported in $[0, \beta(\tau)\tau]$, and it satisfies $\mathbf{m} - \mathbf{m}_\tau = \partial_x \mathbf{M}_\tau$ in $\mathcal{D}'(\mathbf{R})$. Note further that M_τ vanishes as $\tau \rightarrow 0$,

in any $\mathcal{L}^p(\mathbf{R})$ with $1 \leq p < \infty$. Theorem 4.14 suggests that a similar structure can be found in higher space dimensions: the metric projection onto C_ϱ may cause the transfer of momentum to neighboring fluid elements, captured by the distributional divergence $\nabla \cdot \mathbf{M}_\tau$ of the stress tensor field \mathbf{M}_τ . This transfer manifests itself also in the kinetic energy balance, which we will consider next.

Proposition 4.17 (Energy Balance). *For any $(\varrho, \boldsymbol{\mu}, \tau)$ as in Definition 4.9 consider the quantities $(\mathbb{T}_\tau, \mathbb{U}_\tau, \mathbb{W}_\tau, \boldsymbol{\mu}_\tau)$ specified there. Let $\mathbf{M}_\tau \in \mathcal{M}(\beta \mathbf{R}^d; \mathcal{S}_+^d)$ be the matrix field (the stress tensor) satisfying (4.20)/(4.21). Then we have*

$$\begin{aligned} & \int_{\mathbf{R}^{2d}} \frac{1}{2} |\zeta|^2 \boldsymbol{\mu}_\tau(dz, d\zeta) + \int_{\mathbf{R}^d} \text{tr}(\mathbf{M}_\tau(dx)) + \frac{1}{2} A_\tau(\boldsymbol{\mu}, \boldsymbol{\mu}_\tau)^2 \\ &= \int_{\mathbf{R}^{2d}} \frac{1}{2} |\xi|^2 \boldsymbol{\mu}(dx, d\xi), \end{aligned}$$

with acceleration cost (redefined from (3.4))

$$\begin{aligned} A_\tau(\boldsymbol{\mu}, \boldsymbol{\mu}_\tau)^2 &:= \int_{\mathbf{R}^{2d}} \left\{ \frac{3}{4\tau^2} |(x + \tau\xi) - \mathbb{T}_\tau(x, \xi)|^2 \right. \\ &\quad \left. + |\mathbb{W}_\tau(x, \xi) - \mathbb{U}_\tau(x, \xi)|^2 \right\} \boldsymbol{\mu}(dx, d\xi). \end{aligned}$$

Proof. Since $\mathbb{U}_\tau := \mathbb{P}_{\mathcal{H}_\mu(\mathbb{T}_\tau)}(\mathbb{W}_\tau)$ (orthogonal projection), we have

$$\begin{aligned} & \int_{\mathbf{R}^{2d}} |\mathbb{U}_\tau(x, \xi)|^2 \boldsymbol{\mu}(dx, d\xi) + \int_{\mathbf{R}^{2d}} |\mathbb{W}_\tau(x, \xi) - \mathbb{U}_\tau(x, \xi)|^2 \boldsymbol{\mu}(dx, d\xi) \\ &= \int_{\mathbf{R}^{2d}} |\mathbb{W}_\tau(x, \xi)|^2 \boldsymbol{\mu}(dx, d\xi). \end{aligned} \quad (4.23)$$

On the other hand, by definition (4.13) of \mathbb{W}_τ we can write

$$\begin{aligned} & \int_{\mathbf{R}^{2d}} |\mathbb{W}_\tau(x, \xi)|^2 \boldsymbol{\mu}(dx, d\xi) + \frac{3}{4\tau^2} \int_{\mathbf{R}^{2d}} |(x + \tau\xi) - \mathbb{T}_\tau(x, \xi)|^2 \boldsymbol{\mu}(dx, d\xi) \\ &= \int_{\mathbf{R}^{2d}} |v|^2 \boldsymbol{\mu}(dx, d\xi) - \frac{3}{\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathbb{T}_\tau(x, \xi), \mathbb{T}_\tau(x, \xi) \rangle \boldsymbol{\mu}(dx, d\xi) \\ &\quad + \frac{3}{\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathbb{T}_\tau(x, \xi), x \rangle \boldsymbol{\mu}(dx, d\xi) \end{aligned} \quad (4.24)$$

The second integral on the right-hand side of (4.24) vanishes because of (4.4), the last one can be expressed in terms of the stress tensor field \mathbf{M}_τ ; see (4.21). \square

5. ENERGY MINIMIZATION: POLYTROPIC GASES

We now modify the minimization problem of Section 4.2 for polytropic gases. In this case, the density ϱ must be absolutely continuous with respect to the Lebesgue measure since otherwise the internal energy would be infinite (see Definition 5.9). We need a lower semicontinuity result for the internal energy, suitably redefined as a convex functional on the set of monotone transports.

5.1. Gradient Young Measures. We introduce Young measures to capture oscillations and concentrations of weak* converging sequences of derivatives of functions of bounded variations. They will be used in Section 5.2 to establish a lower semicontinuity result for the internal energy. We follow the presentation of [46, 57].

Let $\Omega \subset \mathbf{R}^d$ be a bounded Lipschitz domain and $u \in \text{BV}(\Omega; \mathbf{R}^d)$. Let $\mathcal{B}^{d \times d}$ be the open unit ball in $\mathbf{R}^{d \times d}$ and $\partial \mathcal{B}^{d \times d}$ its boundary. We associate to the derivative Du (which is a measure) a triple $v = (\nu, \sigma, \mu)$ with

$$\nu \in \mathcal{L}_w^\infty(\Omega; \mathcal{M}_{+,1}(\mathbf{R}^{d \times d})), \quad \sigma \in \mathcal{M}_+(\bar{\Omega}), \quad \mu \in \mathcal{L}_w^\infty(\bar{\Omega}, \sigma; \mathcal{M}_{+,1}(\partial \mathcal{B}^{d \times d})) \quad (5.1)$$

(with $\mathcal{M}_{+,1}$ the space of nonnegative σ -additive measures with unit mass) as follows: Consider the Lebesgue-Radon-Nikodým decomposition

$$Du = \nabla u \mathcal{L}^d + D^s u, \quad D^s u \perp \mathcal{L}^d, \quad (5.2)$$

and define $\nu_x := \delta_{\nabla u(x)}$ for a.e. $x \in \Omega$ and $\sigma := |D^s u|$. Let further

$$D^s u = \frac{dD^s u}{d|D^s u|} |D^s u|, \quad p := \frac{dD^s u}{d|D^s u|} \in \mathcal{L}^1(\Omega, |D^s u|; \partial \mathcal{B}^{d \times d}).$$

be the polar decomposition of $D^s u$ and define $\mu_x = \delta_{p(x)}$ for $|D^s u|$ -a.e. $x \in \Omega$. Here $\mathcal{L}_w^\infty(\Omega; \mathcal{M}_{+,1}(\mathbf{R}^{d \times d}))$ is the space of weakly measurable maps from Ω into the space of probability measures on $\mathbf{R}^{d \times d}$ (similar definition for $\mathcal{L}_w^\infty(\bar{\Omega}, \sigma; \mathcal{M}_{+,1}(\partial \mathcal{B}^{d \times d}))$). We call $v = (\nu, \sigma, \mu)$ an elementary Young measure associated to Du .

Consider now a sequence of uniformly bounded maps $u^k \in \text{BV}(\Omega; \mathbf{R}^d)$. Extracting a subsequence if necessary, we may assume that $u^k \rightarrow u$ in $\mathcal{L}^1(\Omega; \mathbf{R}^d)$ and $Du^k \rightharpoonup Du$ weak* in $\mathcal{M}(\Omega; \mathbf{R}^{d \times d})$, for some $u \in \text{BV}(\Omega; \mathbf{R}^d)$. In this case, we say that u^k converges weak* to u in $\text{BV}(\Omega; \mathbf{R}^d)$. We denote by $v^k = (\nu^k, \sigma^k, \mu^k)$ the elementary Young measure associated to Du^k as above. Since the spaces in (5.1) are contained in the dual spaces to $\mathcal{L}^1(\Omega; \mathcal{C}_0(\mathbf{R}^{d \times d}))$, $\mathcal{C}(\bar{\Omega})$, and $\mathcal{L}^1(\bar{\Omega}, \sigma; \mathcal{C}(\partial \mathcal{B}^{d \times d}))$ respectively, one can show that there exists a subsequence (which we do not relabel, for simplicity) and a triple $v = (\nu, \sigma, \mu)$ as in (5.1) with the property that the

$$\begin{aligned} \langle f, v^k \rangle &:= \int_{\Omega} \langle f(x, \cdot), \nu_x^k \rangle dx + \int_{\bar{\Omega}} \langle f^\infty(x, \cdot), \mu_x^k \rangle \sigma^k(dx) \\ &:= \int_{\Omega} \int_{\mathbf{R}^{d \times d}} f(x, A) \nu_x^k(dA) dx + \int_{\bar{\Omega}} \int_{\partial \mathcal{B}^{d \times d}} f^\infty(x, A) \mu_x^k(dA) \sigma^k(dx) \end{aligned} \quad (5.3)$$

converge to $\langle f, v \rangle$ (defined analogously) as $k \rightarrow \infty$, for $f \in \mathbf{R}(\Omega; \mathbf{R}^{d \times d})$ with

$$\mathbf{R}(\Omega; \mathbf{R}^{d \times d}) := \left\{ f: \bar{\Omega} \times \mathbf{R}^{d \times d} \rightarrow \mathbf{R} : \begin{array}{l} \text{the map } f \text{ is a Carathéodory function with} \\ \text{linear growth at infinity, and there exists} \\ f^\infty \in \mathcal{C}(\bar{\Omega} \times \mathbf{R}^{d \times d}) \end{array} \right\}; \quad (5.4)$$

see Corollary 2 and Proposition 2 in [46]. Recall that the map $f: \bar{\Omega} \rightarrow \mathbf{R}^d$ is called a Carathéodory function if it is $\mathcal{L}^d \times \mathcal{B}(\mathbf{R}^{d \times d})$ -measurable and if $A \mapsto f(x, A)$ is continuous for a.e. $x \in \bar{\Omega}$. It is enough to check the measurability of $x \mapsto f(x, A)$ for all $A \in \mathbf{R}^{d \times d}$ fixed; see Proposition 5.6 in [2]. The map f has linear growth at infinity if there exists $M \geq 0$ such that $|f(x, A)| \leq M(1 + \|A\|)$ for a.e. $x \in \bar{\Omega}$ and all $A \in \mathbf{R}^{d \times d}$. We denote by f^∞ the recession function of f , defined as

$$f^\infty(x, A) := \lim_{\substack{x' \rightarrow x \\ A' \rightarrow A \\ t \rightarrow \infty}} \frac{f(x', tA')}{t} \quad \text{for a.e. } x \in \bar{\Omega} \text{ and all } A \in \mathbf{R}^{d \times d}. \quad (5.5)$$

Note that the recession function is positively 1-homogeneous in A , if it exists. We call a triple $v = (\nu, \sigma, \mu)$ obtained as a limit as above a gradient Young measure

and denote the space of gradient Young measures by $\mathbf{GY}(\Omega; \mathbf{R}^{d \times d})$. Then

$$Du = \langle \text{id}, \nu \rangle \mathcal{L}^d + \langle \text{id}, \mu \rangle \sigma, \quad (5.6)$$

by construction (cf. (5.3)). Moreover, we have

$$\|\nabla u^k\| \mathcal{L}^d + \left\| \frac{dD^s u^k}{d|D^s u^k|} \right\| |D^s u^k| \longrightarrow \langle \|\cdot\|, \nu \rangle \mathcal{L}^d + \langle \|\cdot\|, \mu \rangle \sigma$$

weak* in $\mathcal{M}(\bar{\Omega})$ as $k \rightarrow \infty$, which implies that $\langle \|\cdot\|, \nu \rangle \in \mathcal{L}^1(\Omega)$. We used the fact that the recession function of $f(x, A) := \varphi(x)\|A\|$ with $\varphi \in \mathcal{C}(\bar{\Omega})$ coincides with f . We refer the reader to [46] for further information.

We apply this framework to sequences of *monotone* functions $u^k \in \text{BV}(\Omega; \mathbf{R}^d)$ (see Remark 3.1), in which case the derivatives Du^k are positive (that is, matrix-valued and locally finite) measures; see Theorem 5.3 in [1]. We define

$$\mathcal{P}_{++}^d := \left\{ A \in \mathbf{R}^{d \times d} : \langle v, Av \rangle > 0 \text{ for all } v \in \mathbf{R}^d, v \neq 0 \right\};$$

the analogous set with $\langle v, Av \rangle \geq 0$ for $v \in \mathbf{R}^d, v \neq 0$, will be denoted by \mathcal{P}_+^d . Since the map $(A, v) \mapsto \langle v, Av \rangle$ is continuous, the set \mathcal{P}_{++}^d is open and convex, the set \mathcal{P}_+^d is a closed convex cone. Notice that a matrix A is an element of \mathcal{P}_+^d (resp. \mathcal{P}_{++}^d) if and only if its symmetric part $A^{\text{sym}} := (A + A^T)/2 \in \mathcal{S}_+^d$ (resp. \mathcal{S}_{++}^d).

Proposition 5.1 (Gradient Young Measures). *Let $\Omega \subset \mathbf{R}^d$ be a bounded Lipschitz domain and suppose that $u^k \rightharpoonup u$ weak* in $\text{BV}(\Omega; \mathbf{R}^d)$ with $u^k, u \in \text{BV}(\Omega; \mathbf{R}^d)$ monotone. For all $k \in \mathbf{N}$ we denote by v^k the elementary gradient Young measure associated to Du^k , as introduced above. Then there exists a subsequence (which we do not relabel, for simplicity) and a gradient Young measure $v \in \mathbf{GY}(\Omega; \mathbf{R}^{d \times d})$ with the property that $\langle f, v^k \rangle \rightarrow \langle f, v \rangle$ for all $f \in \mathbf{R}_+(\Omega; \mathbf{R}^{d \times d})$, where*

$$\mathbf{R}_+(\Omega; \mathbf{R}^{d \times d}) := \left\{ f : \bar{\Omega} \times \mathbf{R}^{d \times d} \rightarrow \mathbf{R} : \begin{array}{l} \text{the map } f \text{ is a Carathéodory function with} \\ \text{linear growth at infinity, and there exists} \\ f^\infty \in \mathcal{C}(\bar{\Omega} \times \mathcal{P}_+^d) \end{array} \right\}.$$

Proof. It suffices to check continuity of the recession function f^∞ on the smaller set \mathcal{P}_+^d because all gradient Young measures considered above vanish outside of $\Omega \times \mathcal{P}_+^d$. Indeed, consider any test function $f \in \mathbf{R}(\Omega; \mathbf{R}^{d \times d})$ of the form $f(x, A) = \varphi(x)h(A)$, with $\varphi \in \mathcal{C}_c(\Omega)$ nonnegative and $h(A) := \text{dist}(A, \mathcal{P}_+^d)$ for all $A \in \mathbf{R}^{d \times d}$. Then the map h is positively 1-homogeneous. This follows immediately from the fact that \mathcal{P}_+^d is a cone. It can also be derived from the following observation: Notice first that symmetric and antisymmetric matrices in $\mathbf{R}^{d \times d}$ are orthogonal to each other with respect to the Frobenius inner product. For given $A \in \mathbf{R}^{d \times d}$ let $B := (A + A^T)/2$ and $C := (A - A^T)/2$ be the symmetric and antisymmetric parts of A , respectively. Let $B = UH$ be a polar decomposition ($U^T U = \mathbb{1}$, $H = H^T \geq 0$). Then

$$X_A := C + (B + H)/2$$

is the unique element in \mathcal{P}_+^d closest to A in the Frobenius norm, and

$$\text{dist}(A, \mathcal{P}_+^d)^2 = \sum_{\lambda_i(B) < 0} \lambda_i(B)^2,$$

with $\lambda_i(B)$ the (real) eigenvalues of B ; see [43]. Then the claim follows.

Since h is positively 1-homogeneous it is sufficient to consider the limits $x' \rightarrow x$ and $A' \rightarrow A$ in (5.5) to define the recession function of f . But φ, h are continuous, and so f^∞ coincides with f . In particular, this proves that $f \in \mathbf{R}(\Omega; \mathbf{R}^{d \times d})$. Notice that $h(A) = 0$ if and only if $A \in \mathcal{P}_+^d$. If u^k is a monotone map, then $\nabla u^k(x) \in \mathcal{P}_+^d$ for a.e. $x \in \Omega$ and $p^k(x) \in \mathcal{P}_+^d$ for $|D^s u^k|$ -a.e. $x \in \Omega$, where

$$D^s u^k = \frac{dD^s u^k}{d|D^s u^k|} |D^s u^k|, \quad p^k := \frac{dD^s u^k}{d|D^s u^k|} \in \mathcal{L}^1(\Omega, |D^s u^k|; \partial \mathcal{B}^{d \times d})$$

is the polar decomposition of $D^s u^k$. If v^k is the elementary gradient Young measure associated to Du^k , then $\langle f, v^k \rangle = 0$ for all $k \in \mathbf{N}$ and f as above. Then the gradient Young measure v generated by $\{v^k\}$ satisfies $\langle f, v \rangle = 0$ because $\langle f, v^k \rangle \rightarrow \langle f, v \rangle$ for all $f \in \mathbf{R}(\Omega; \mathbf{R}^{d \times d})$. Since $\varphi \in \mathcal{C}_c(\Omega)$ nonnegative was arbitrary, we obtain that the gradient Young measure v vanishes outside of $\Omega \times \mathcal{P}_+^d$ as well. Then a careful inspection of the proof of Proposition 2 in [46] yields the result: The convergence of the gradient Young measures follows from the weak* convergence of the product measures $\nu^k \mathcal{L}^d + \mu^k \sigma^k \rightharpoonup \nu \mathcal{L}^d + \mu \sigma$ on (a suitable compactification of) $\Omega \times \mathbf{R}^{d \times d}$, which reduces to weak* convergence on $\Omega \times \mathcal{P}_+^d$ when u^k and u are monotone. \square

5.2. Internal Energy. We introduce a functional on the space of monotone BV-vector fields that represents the internal energy. This functional will be convex and lower semicontinuous with respect to weak* convergence in $\text{BV}_{\text{loc}}(\Omega; \mathbf{R}^d)$.

Let us start with two auxiliary results.

Lemma 5.2. *For any $\gamma > 1$, the map $h: \mathbf{R}^{d \times d} \rightarrow [0, \infty]$ defined by*

$$h(A) := \begin{cases} \det(A^{\text{sym}})^{1-\gamma} & \text{if } A \in \mathcal{P}_{++}^d, \\ \infty & \text{otherwise,} \end{cases} \quad (5.7)$$

is lower semicontinuous, proper, and convex. For all $A \in \mathbf{R}^{d \times d}$, we have

$$h^\infty(A) := \lim_{t \rightarrow \infty} \frac{h(\mathbb{1} + tA) - h(\mathbb{1})}{t} = \begin{cases} 0 & \text{if } A \in \mathcal{P}_+^d, \\ \infty & \text{otherwise.} \end{cases} \quad (5.8)$$

Proof. Since $A \mapsto \det(A^{\text{sym}})$ is continuous, the function h is lower semicontinuous. It is proper because $h(\mathbb{1}) = 1$. In order to prove the convexity of h , we observe that $A \mapsto \det(A)^{1/d}$ is concave for all symmetric, positive definite $A \in \mathbf{R}^{d \times d}$. Indeed, pick any two such matrices A^0 and A^1 . For all $s \in [0, 1]$ we can write

$$\det((1-s)A^0 + sA^1)^{1/d} = (\det(A^0) \det(\mathbb{1} + sB))^{1/d},$$

where $B := C^{-1}(A^1 - A^0)C^{-1}$ and $C := \sqrt{A^0}$. The matrix C exists and is invertible since A^0 is symmetric and positive definite, by assumption. Then we compute

$$\begin{aligned} \frac{d}{ds} \det(\mathbb{1} + sB)^{1/d} &= \det(\mathbb{1} + sB)^{1/d} \left\{ \frac{1}{d} \text{tr}(D) \right\}, \\ \frac{d^2}{ds^2} \det(\mathbb{1} + sB)^{1/d} &= \det(\mathbb{1} + sB)^{1/d} \left\{ \frac{1}{d^2} \text{tr}(D)^2 - \frac{1}{d} \text{tr}(D^2) \right\}, \end{aligned} \quad (5.9)$$

where $D := B(\mathbb{1} + sB)^{-1}$. The matrix D is symmetric. Therefore

$$\text{tr}(D)^2 = (\lambda_1 + \dots + \lambda_d)^2 \leq d(\lambda_1^2 + \dots + \lambda_d^2) = d \text{tr}(D^2),$$

where $\lambda_1, \dots, \lambda_d$ are the real eigenvalues of D . Hence (5.9) is nonpositive for every $s \in [0, 1]$. The composition of a concave function with a convex, nonincreasing map

is convex. Therefore the map $A \mapsto \det(A)^{1-\gamma}$ is convex for all symmetric, positive definite $A \in \mathbf{R}^{d \times d}$. Finally, the composition of any convex function with the linear map $A \mapsto A^{\text{sym}}$ is again convex. Then the result follows.

To prove (5.8), we use that the map $t \mapsto (h(\mathbb{1} + tA) - h(\mathbb{1}))/t$ is nondecreasing (and hence $\lim_{t \rightarrow \infty} = \sup_{t > 0}$), by convexity of h . If now $A \notin \mathcal{P}_+^d$, then there exists $v \in \mathbf{R}^d$, $\|v\| = 1$, such that $\langle v, Av \rangle < 0$. For sufficiently large $t > 0$, we get

$$\langle v, (\mathbb{1} + tA)v \rangle = 1 + t\langle v, Av \rangle < 0,$$

and thus $h(\mathbb{1} + tA) = \infty$. This proves (5.8) for $A \notin \mathcal{P}_+^d$.

On the other hand, if $A \in \mathcal{P}_+^d$, then $\mathbb{1} + tA \in \mathcal{P}_{++}^d$ for all $t > 0$, because

$$\langle v, (\mathbb{1} + tA)v \rangle = 1 + t\langle v, Av \rangle \geq 1$$

for all $v \in \mathbf{R}^d$, $\|v\| = 1$. By convexity of $A \mapsto \det(A^{\text{sym}})^{1/d}$, we obtain

$$\begin{aligned} \det(\mathbb{1} + tA^{\text{sym}})^{1/d} &= \det \left((1+t) \left(\frac{1}{1+t} \mathbb{1} + \frac{t}{1+t} A^{\text{sym}} \right) \right)^{1/d} \\ &\geq (1+t) \left(\frac{1}{1+t} \det(\mathbb{1})^{1/d} + \frac{t}{1+t} \det(A^{\text{sym}})^{1/d} \right) \geq 1 \end{aligned}$$

for all $t > 0$. Notice that $\det(A^{\text{sym}}) \geq 0$. This implies that $\det(\mathbb{1} + tA^{\text{sym}})^{1-\gamma} \leq 1$ (recall that $\gamma > 1$, by assumption), and so (5.8) follows for $A \in \mathcal{P}_+^d$ as well. \square

Lemma 5.3. *For any $n \in \mathbf{N}$, we define the inf-convolution*

$$h_n(A) := \inf_{B \in \mathbf{R}^{d \times d}} \{n\|A - B\| + h(B)\} \quad (5.10)$$

for all $A \in \mathbf{R}^{d \times d}$, which has the following properties:

- (1) The map h_n is lower semicontinuous, proper, and convex.
- (2) For all $A \in \mathbf{R}^{d \times d}$, we have $h_n(A) \rightarrow h(A)$ monotonically from below.
- (3) The map h_n is Lipschitz continuous with Lipschitz constant n .
- (4) The map h_n has linear growth at infinity:

$$h_n(A) \leq 1 + n\sqrt{d} + n\|A\| \quad \text{for all } A \in \mathbf{R}^{d \times d}. \quad (5.11)$$

- (5) For all $A \in \mathbf{R}^{d \times d}$, we have that

$$h_n^\infty(A) := \lim_{t \rightarrow \infty} \frac{h_n(\mathbb{1} + tA) - h_n(\mathbb{1})}{t} = n \text{dist}(A, \mathcal{P}_+^d). \quad (5.12)$$

Proof. Statement (1) follows from Corollary 9.2.2 in [59]: Notice first that the norm and h are lower semicontinuous, convex, and proper; see Lemma 5.2. The recession function of the norm is the norm itself, and it holds

$$n\|A\| + h^\infty(-A) > 0 \quad \text{for all } A \in \mathbf{R}^{d \times d}, A \neq 0.$$

Statements (2) and (3) follow from Lemma 1.61 in [2].

To prove (5.11), we just choose $B = \mathbb{1}$ in (5.10) and use the triangle inequality.

Finally, statement (5) follows from Corollary 9.2.1 in [59]. We must prove that for all pairs of matrices $A_1, A_2 \in \mathbf{R}^{d \times d}$ with the property that

$$n\|A_1\| + h^\infty(A_2) \leq 0 \quad \text{and} \quad n\|A_1\| + h^\infty(-A_2) > 0, \quad (5.13)$$

it holds $A_1 + A_2 \neq 0$. The first condition in (5.13) is only satisfied if $A_1 = 0$ and $A_2 \in \mathcal{P}_+^d$, because of (5.8). Then the second condition requires $-A_2 \notin \mathcal{P}_+^d$. That means, there exists $v \in \mathbf{R}^d$, $v \neq 0$, with $\langle v, -A_2 v \rangle < 0$ (this is consistent with

$A_2 \in \mathcal{P}_+^d$). This is only possible if $A_2 \neq 0$, so the claim follows. We then obtain that the recession function (5.12) of the inf-convolution (5.10) is given by

$$h_n^\infty(A) = \inf_{B \in \mathbf{R}^{d \times d}} \left\{ n\|A - B\| + h^\infty(B) \right\}$$

for all $A \in \mathbf{R}^{d \times d}$, which implies the result because of (5.8). \square

We can now prove the following lower semicontinuity result.

Proposition 5.4 (Internal Energy). *Let $\Omega \subset \mathbf{R}^d$ be open and convex, and h given by (5.7). For $e \in \mathcal{L}^1(\Omega)$ nonnegative and $u \in \text{BV}_{\text{loc}}(\Omega; \mathbf{R}^d)$ we define*

$$\mathcal{U}[u] := \begin{cases} \int_{\Omega} e(x)h(\nabla u(x)) \, dx & \text{if } u \text{ monotone,} \\ \infty & \text{otherwise,} \end{cases} \quad (5.14)$$

using again the decomposition (5.2). Then the following is true:

- (1) The functional \mathcal{U} is convex.
- (2) For any $u^k \rightharpoonup u$ weak* in $\text{BV}_{\text{loc}}(\Omega; \mathbf{R}^d)$ with $u^k, u \in \text{BV}_{\text{loc}}(\Omega; \mathbf{R}^d)$ monotone, there exists a subsequence (not relabeled) such that

$$\mathcal{U}[u] \leq \liminf_{k \rightarrow \infty} \mathcal{U}[u^k].$$

Remark 5.5. Notice that in (5.14) we only consider the part of Du that is absolutely continuous with respect to \mathcal{L}^d and disregard the singular component. The intuition is that (for each direction) only increasing jumps are allowed in the transport map u , which correspond to the formation of vacuum, which is admissible.

Proof of Proposition 5.4. We proceed in two steps.

Step 1. Consider $u^k \in \text{BV}_{\text{loc}}(\Omega; \mathbf{R}^d)$ with $k = 0..1$. For any $s \in (0, 1)$ we define $u^s := (1-s)u^0 + su^1 \in \text{BV}_{\text{loc}}(\Omega; \mathbf{R}^d)$. If $\mathcal{U}[u^k] = \infty$ for $k = 0$ or $k = 1$, then there is nothing to prove, so we may assume that both terms are finite. This requires that both u^k are monotone and $\nabla u^k(x) \in \mathcal{P}_{++}^d$ for $e\mathcal{L}^d$ -a.e. $x \in \Omega$. It follows that u^s is monotone as well and $\nabla u^s(x) \in \mathcal{P}_{++}^d$ for $e\mathcal{L}^d$ -a.e. $x \in \Omega$. Then

$$h(\nabla u^s(x)) \leq (1-s)h(\nabla u^0(x)) + sh(\nabla u^1(x));$$

see Lemma 5.2. Multiplying by $e(x)$ and integrating in Ω , we obtain

$$\mathcal{U}[u^s] \leq (1-s)\mathcal{U}[u^0] + s\mathcal{U}[u^1]$$

for all $s \in [0, 1]$. This proves the convexity of the functional.

Step 2. We first introduce a sequence of bounded open convex sets

$$\Omega_n := \left\{ x \in B_n(0) : \text{dist}(x, \mathbf{R}^d \setminus \Omega) > 1/n \right\},$$

which are bounded Lipschitz domains. We have $\Omega_{n-1} \subset \Omega_n$ for all $n \in \mathbf{N}$.

We then choose a sequence of cut-off functions $\varphi_n \in \mathcal{C}_c(\Omega; [0, 1])$ with $\varphi_n(x) = 1$ for all $x \in \Omega_{n-1}$ and $\varphi_n(x) = 0$ for all $x \notin \Omega_n$. For all $n \in \mathbf{N}$ we define

$$f_n(x, A) := (e(x) \wedge n)\varphi_n(x)h_n(A) \quad \text{for all } (x, A) \in \Omega \times \mathbf{R}^{d \times d}, \quad (5.15)$$

where h_n is given by (5.10). Because of Lemma 5.3, the map f_n is a Carathéodory function with linear growth at infinity. In fact, we can estimate

$$0 \leq f_n(x, A) \leq n(1 + n\sqrt{d} + n\|A\|) \quad \text{for all } (x, A) \in \Omega \times \mathbf{R}^{d \times d}.$$

We prove that $f_n^\infty(x, A) = 0$ for all $(x, A) \in \Omega \times \mathcal{P}_+^d$: Note first that

$$\left| \frac{f_n(x', tA')}{t} - 0 \right| \leq nh_n(tA')/t \leq n \left\{ h_n(tA)/t + n\|A' - A\| \right\},$$

uniformly in $x' \in \Omega$. Recall that h_n is Lipschitz continuous with Lipschitz constant n . Since $h_n(A) < \infty$ for all $A \in \mathbf{R}^{d \times d}$, by Theorem 8.5 in [59] we have

$$h_n(tA)/t \longrightarrow h_n^\infty(A) \quad \text{as } t \rightarrow \infty,$$

which vanishes for any $A \in \mathcal{P}_+^d$; see Lemma 5.3. This implies that $f_n^\infty \in \mathcal{C}(\Omega \times \mathcal{P}_+^d)$, and so $f_n \in \mathbf{R}_+(\Omega_n; \mathbf{R}^{d \times d})$ for all $n \in \mathbf{N}$. By construction, it holds

$$f_n(x, A) \leq f_{n+1}(x, A) \quad \text{and} \quad e(x)h(A) = \sup_n f_n(x, A) \quad (5.16)$$

for all $(x, A) \in \Omega \times \mathbf{R}^{d \times d}$. We used again Lemma 5.3.

Let us fix an $n \in \mathbf{N}$ for a moment. Extracting a subsequence if necessary, we may assume that the sequence of elementary gradient Young measures v^k generated by $Du^k|_{\Omega_n}$ converges to $v = (\nu, \sigma, \mu) \in \mathbf{GY}(\Omega_n; \mathbf{R}^{d \times d})$ in the sense that

$$\langle\langle f, v^k \rangle\rangle \longrightarrow \langle\langle f, v \rangle\rangle \quad \text{for all } f \in \mathbf{R}_+(\Omega_n; \mathbf{R}^{d \times d}); \quad (5.17)$$

see Proposition 5.1. It holds $Du = \langle \text{id}, \nu \rangle \mathcal{L}^d + \langle \text{id}, \mu \rangle \sigma$. Comparing this identity with the Lebesgue-Radon-Nikodým decomposition (5.2), we find

$$\nabla u = \langle \text{id}, \nu \rangle + \langle \text{id}, \mu \rangle \frac{d\sigma}{d\mathcal{L}^d} \quad \text{a.e.} \quad \text{and} \quad D^s u = \langle \text{id}, \mu \rangle \sigma^s, \quad (5.18)$$

where $\sigma^s \perp \mathcal{L}^d$ is the singular part of σ . Notice that $\langle \text{id}, \mu_x \rangle \in \mathbf{R}^{d \times d}$ may not have unit length for σ^s -a.e. $x \in \Omega$. The polar decomposition of $D^s u$ is given by

$$|D^s u| = |\langle \text{id}, \mu \rangle| \sigma^s \quad \text{and} \quad \frac{dD^s u}{d|D^s u|} = \frac{\langle \text{id}, \mu \rangle}{|\langle \text{id}, \mu \rangle|} \quad |D^s u| \text{-a.e.}$$

We apply the convergence (5.17) to the function f_n defined in (5.15), whose restriction to Ω_n belongs to $\mathbf{R}_+(\Omega_n; \mathbf{R}^{d \times d})$. We observe that

$$f_n^\infty(\cdot, \langle \text{id}, \mu \rangle) \sigma^s = f_n^\infty\left(\cdot, \frac{dD^s u}{d|D^s u|}\right) |D^s u|$$

because the map $A \mapsto f_n^\infty(x, A)$ is positively 1-homogeneous for $x \in \Omega_n$. Then the following Jensen-type inequalities hold (see Theorem 9 in [46]):

$$f_n(\cdot, \nabla u) \leq \langle f_n, \nu \rangle + \langle f_n^\infty, \mu \rangle \frac{d\sigma}{d\mathcal{L}^d} \quad \text{a.e.},$$

$$f_n^\infty(\cdot, \langle \text{id}, \mu \rangle) \leq \langle f_n^\infty, \mu \rangle \quad \sigma^s \text{-a.e.}$$

because the map $A \mapsto f_n(x, A)$ is convex for $x \in \Omega$. We can then estimate

$$\begin{aligned} & \lim_{k \rightarrow \infty} \int_{\Omega_n} f_n(\cdot, \nabla u^k) + \int_{\Omega_n} f_n^\infty\left(\cdot, \frac{dD^s u^k}{d|D^s u^k|}\right) |D^s u^k| \\ &= \int_{\Omega_n} \langle f_n, \nu \rangle + \int_{\Omega_n} \langle f_n^\infty, \mu \rangle \sigma \\ &= \int_{\Omega_n} \left(\langle f_n, \nu \rangle + \langle f_n^\infty, \mu \rangle \frac{d\sigma}{d\mathcal{L}^d} \right) + \int_{\Omega_n} \langle f_n^\infty, \mu \rangle \sigma^s \\ &\geq \int_{\Omega_n} f_n(\cdot, \nabla u) + \int_{\Omega_n} f_n^\infty\left(\cdot, \frac{dD^s u}{d|D^s u|}\right) |D^s u|. \end{aligned}$$

Clearly the integrals can be extended to all of Ω because f_n vanishes outside of Ω_n . Moreover, we have shown that the recession function f_n^∞ vanishes. Hence

$$\int_{\Omega} f_n(x, \nabla u(x)) dx \leq \liminf_{k \rightarrow \infty} \int_{\Omega} e(x) h(\nabla u^k(x)) dx, \quad (5.19)$$

where we also used (5.16). By a standard diagonal argument (successively extracting subsequences if necessary), we may assume that (5.19) holds for all $n \in \mathbf{N}$. We then use (5.16) and the monotone convergence theorem to obtain the result. \square

We finish the section with an estimate on determinants of square matrices.

Lemma 5.6. *Suppose S is a real, positive semidefinite, symmetric $(d \times d)$ -matrix. For any real skew-symmetric $(d \times d)$ -matrix A we have*

$$\det(S + A) \geq \det S \geq 0. \quad (5.20)$$

Proof. We divide the proof into two steps.

Step 1. We will first prove that if $\det S = 0$, then $\det(S + A) \geq 0$. Recall that the determinants of square matrices equal the product of their eigenvalues. Non-real eigenvalues of $S + A$ can only occur in complex conjugate pairs because S, A are real matrices. Since the product of two complex conjugate numbers is nonnegative, it remains to prove that every *real* eigenvalue of $S + A$ must be nonnegative. Let $\lambda \in \mathbf{R}$ be an eigenvalue with corresponding eigenvector v . Note that if v is complex, then its complex conjugate \bar{v} is another eigenvector to the same eigenvalue λ . Taking the sum $v + \bar{v}$ if necessary, we may therefore assume that $v \in \mathbf{R}^d$. We have

$$(S + A)v = \lambda v, \quad \|v\| > 0.$$

We take the inner product with v and obtain (since A is skew-symmetric)

$$\lambda \|v\|^2 = \langle (S + A)v, v \rangle = \langle Sv, v \rangle.$$

The right-hand side is nonnegative because S is positive semidefinite. Hence $\lambda \geq 0$. From this, we conclude that $\det(S + A) \geq \det S$ whenever $\det S = 0$.

Step 2. Consider now $\det S \neq 0$. Since S is positive semidefinite and symmetric, all eigenvalues of S (which are real) are positive. Therefore $\det S > 0$ and $\langle Sv, v \rangle > 0$ for every $v \in \mathbf{R}^d$ with $v \neq 0$. We claim that $\det(S + tA) > 0$ for every $t \in \mathbf{R}$. In fact, assume this is false. Then zero is an eigenvalue of $S + tA$, with corresponding eigenvector $v \in \mathbf{R}^d$ (see above). We have $(S + tA)v = 0$ and $v \neq 0$. We get

$$0 < \langle Sv, v \rangle = \langle (S + tA)v, v \rangle = 0,$$

using again that A is skew-symmetric. This contradiction proves the claim.

For all $t \in \mathbf{R}$, we can now define $f(t) := \log \det(S + tA)$. We compute

$$f'(t) = \text{tr}((S + tA)^{-1}A).$$

Notice that $t(S + tA)^{-1}A = \text{id} - (S + tA)^{-1}S$. Since S is symmetric, there exists an orthogonal matrix P such that $P^{-1}SP = \Lambda$, where $\Lambda := \text{diag}(\lambda_1, \dots, \lambda_d)$ contains the eigenvalues $\lambda_i > 0$ of S , $i = 1 \dots d$. Let e_i denote the i th standard basis vector of \mathbf{R}^d . Since the trace is invariant under changes of basis, we obtain

$$\begin{aligned} \text{tr}((S + tA)^{-1}A) &= \text{tr}(\text{id} - P^{-1}(S + tA)^{-1}SP) \\ &= \sum_{i=1}^d (1 - \langle P^{-1}(S + tA)^{-1}SP e_i, e_i \rangle). \end{aligned}$$

We denote by v_i the i th column vector of P (hence $v_i = Pe_i$), which is a normalized eigenvector of S corresponding to the eigenvalue λ_i . As $P^{-1} = P^T$, we have

$$\text{tr}((S + tA)^{-1}A) = \sum_{i=1}^d (1 - \lambda_i \langle w_i, v_i \rangle), \quad w_i := (S + tA)^{-1}v_i. \quad (5.21)$$

Since the eigenvectors v_1, \dots, v_d form an orthonormal basis of \mathbf{R}^d , there is a unique expansion $w_i = \sum_{k=1}^d \alpha_i^k v_k$ with $\alpha_i^k := \langle w_i, v_k \rangle$. Using this expansion, we get

$$\alpha_i^i = \langle w_i, (S + tA)w_i \rangle = \langle w_i, Sw_i \rangle = \sum_{k=1}^d \lambda_k (\alpha_i^k)^2 \quad (5.22)$$

for $i = 1 \dots d$. Recall that the eigenvalues λ_k are all positive and A is skew-symmetric. We conclude that $\alpha_i^i \geq 0$. Moreover, rewriting (5.22) in the form

$$\alpha_i^i (1 - \lambda_i \alpha_i^i) = \sum_{k \neq i} \lambda_k (\alpha_i^k)^2 \geq 0,$$

we obtain that $1 - \lambda_i \langle w_i, v_i \rangle \geq 0$ for each i . Using this estimate in (5.21), we conclude that $f'(t) \geq 0$ for all $t > 0$, and so the map $t \mapsto f(t)$ is nondecreasing for such t . In particular, we have that $\det(S + A) = \exp f(1) \geq \exp f(0) = \det S > 0$. \square

5.3. Minimization Problem. We now introduce the main minimization problem for (1.1). We represent the state of the fluid by $(\varrho, \boldsymbol{\mu}, \sigma)$, with $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ the density, $\boldsymbol{\mu} \in \mathcal{P}_2(\mathbf{R}^{2d})$ the velocity distribution, and $\sigma \in \mathcal{M}_+(\mathbf{R}^d)$ the thermodynamic entropy. We assume that $\mathcal{U}[\varrho, \sigma] < \infty$, which implies that $\varrho = r\mathcal{L}^d$ and $\sigma = \varrho S$ for suitable Borel functions r, S ; see Definition 1.1. In the isentropic case, S will be constant in time and space. We want to minimize the sum of the internal energy of the transported fluid and the acceleration cost of the transport, over the cone $\mathcal{C}_{\boldsymbol{\mu}}$ of monotone maps; see Definition 4.1. The following observation will be useful:

Lemma 5.7. *Let $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ and $\boldsymbol{\mu} \in \mathcal{P}_2(\mathbf{R}^{2d})$, where $\varrho \ll \mathcal{L}^d$. To every $\mathbb{T} \in \mathcal{C}_{\boldsymbol{\mu}}$ we can associate a function $\mathcal{T} \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$ defined on all of \mathbf{R}^d that satisfies*

$$\mathbb{T}(x, \xi) = \mathcal{T}(x) \quad \text{for } \boldsymbol{\mu}\text{-a.e. } (x, \xi) \in \mathbf{R}^{2d}. \quad (5.23)$$

The map \mathcal{T} is monotone on $\Omega := \text{int } \overline{\text{conv spt } \varrho}$ (hence $\mathcal{T} \in \text{BV}_{\text{loc}}(\Omega; \mathbf{R}^d)$):

$$\langle \mathcal{T}(x_1) - \mathcal{T}(x_2), x_1 - x_2 \rangle \geq 0 \quad \text{for all } x_1, x_2 \in \Omega.$$

Proof. For $\mathbb{T} \in \mathcal{C}_{\boldsymbol{\mu}}$ let u be any maximal monotone map associated to $\gamma := (\mathbb{x}, \mathbb{T}) \# \boldsymbol{\mu}$, which is in C_ϱ ; see Definition 3.3. As shown in Lemma 3.4, the domain of u contains the convex open set Ω . As $\varrho \ll \mathcal{L}^d$, the set Ω must be nonempty and $\varrho(\mathbf{R}^d \setminus \Omega) = 0$ (since the boundary of $\overline{\text{conv spt } \varrho}$ is a Lipschitz manifold of codimension one, which is a Lebesgue null set and hence ϱ -negligible). Consequently, the maximal monotone map u associated to γ is defined ϱ -a.e. The map u is single-valued for all $x \in \Omega \setminus \Sigma^1(u)$ (see Remark 3.1), and $\Sigma^1(u)$ is a Lebesgue null set and hence ϱ -negligible. We now define a (single-valued) function \mathcal{T} on all of \mathbf{R}^d as follows:

$$\mathcal{T}(x) := \begin{cases} v & \text{if } x \in \Omega \setminus \Sigma^1(u) \text{ and } u(x) =: \{v\}, \\ \bar{v} & \text{if } x \in \Omega \cap \Sigma^1(u) \text{ and } \bar{v} \text{ is the center of mass of } u(x), \\ 0 & \text{if } x \in \mathbf{R}^d \setminus \Omega. \end{cases}$$

Then \mathcal{T} is monotone on Ω because $\mathcal{T}(x) \in u(x)$ for every $x \in \Omega$. Recall that $u(x)$ is a closed convex set (possibly empty) for all $x \in \mathbf{R}^d$; see Proposition 1.2 in [1].

As shown in Remark 4.4, there exists a Borel set $N_{\mathbb{T}} \subset \mathbf{R}^{2d}$ such that $\mu(N_{\mathbb{T}}) = 0$ and $(x, \mathbb{T}(x, \xi)) \in \text{spt } \gamma$ for all $(x, \xi) \in \mathbf{R}^{2d} \setminus N_{\mathbb{T}}$. This implies that $\mathbb{T}(x, \xi) \in u(x)$ for such (x, ξ) , since $\text{graph}(u)$ is an extension of $\text{spt } \gamma$. Therefore

$$\begin{aligned} \left\{ (x, \xi) \in \mathbf{R}^{2d} : \mathbb{T}(x, \xi) \neq \mathcal{T}(x) \right\} &\subset N_{\mathbb{T}} \cup (E \times \mathbf{R}^d), \\ \text{where } E &:= (\mathbf{R}^d \setminus \Omega) \cup (\Omega \cap \Sigma^1(u)). \end{aligned}$$

Since $\mu(N_{\mathbb{T}}) = 0$ and $\mu(E \times \mathbf{R}^d) = \varrho(E) = 0$, statement (5.23) follows. Now

$$\int_{\mathbf{R}^d} |\mathcal{T}(x)|^2 \varrho(dx) = \int_{\mathbf{R}^{2d}} |\mathcal{T}(x)|^2 \mu(dx, d\xi) = \int_{\mathbf{R}^{2d}} |\mathbb{T}(x, \xi)|^2 \mu(dx, d\xi),$$

which is finite. For $\mathcal{T} \in \text{BV}_{\text{loc}}(\Omega; \mathbf{R}^d)$ we refer the reader to Theorem 5.3 in [1]. \square

Lemma 5.7 shows that instead of minimizing over \mathcal{C}_{μ} it is sufficient to consider a minimization over the following convex cone in $\mathcal{L}^2(\mathbf{R}^d, \varrho)$ (we refer the reader to the proof of Proposition 5.13 for topological properties):

Definition 5.8 (Configurations). Let $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$ satisfy $\varrho \ll \mathcal{L}^d$. We denote by \mathcal{C}_{ϱ} the set of all Borel maps $\mathcal{T} : \mathbf{R}^d \rightarrow \mathbf{R}^d$ with the following properties:

- (1) \mathcal{T} is monotone on $\Omega := \text{int } \overline{\text{conv spt } \varrho}$ (hence $\mathcal{T} \in \text{BV}_{\text{loc}}(\Omega; \mathbf{R}^d)$),
- (2) $\mathcal{T} \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$.

If $\mu \in \mathcal{P}_{\varrho}(\mathbf{R}^{2d})$ and $\mathbb{T} \in \mathcal{C}_{\mu}$ are given, and $\tau > 0$, then

$$\begin{aligned} \int_{\mathbf{R}^{2d}} |(x + \tau\xi) - \mathbb{T}(x, \xi)|^2 \mu(dx, d\xi) \\ = \tau^2 \int_{\mathbf{R}^{2d}} |\xi - \mathbf{u}(x)|^2 \mu(dx, d\xi) + \int_{\mathbf{R}^d} |(x + \tau\mathbf{u}(x)) - \mathcal{T}(x)|^2 \varrho(dx), \end{aligned} \quad (5.24)$$

for every map $\mathcal{T} \in \mathcal{C}_{\varrho}$ satisfying (5.23). Here \mathbf{u} is the barycentric projection $\mathbb{B}(\mu)$ of μ (equivalently, the orthogonal projection of μ onto the space of functions in $\mathcal{L}^2(\mathbf{R}^{2d}, \mu)$ that depend only on the spatial variable $x \in \mathbf{R}^d$). Notice that the first term on the right-hand side of (5.24) does not depend on \mathbb{T} or \mathcal{T} .

For any smooth, strictly monotone map $\mathcal{T} : \mathbf{R}^d \rightarrow \mathbf{R}^d$, the internal energy of the fluid transported by \mathcal{T} is given (after a change of variables) by

$$\begin{aligned} \mathcal{U}[\mathcal{T} \# \varrho, \mathcal{T} \# \sigma] &= \int_{\mathbf{R}^d} U \left(\left(\frac{r}{\det(\nabla \mathcal{T})} \right) \circ \mathcal{T}^{-1}(z), S \circ \mathcal{T}^{-1}(z) \right) dz \\ &= \int_{\mathbf{R}^d} U \left(\frac{r(x)}{\det(\nabla \mathcal{T}(x))}, S(x) \right) \det(\nabla \mathcal{T}(x)) dx \\ &= \int_{\mathbf{R}^d} U(r(x), S(x)) \det(\nabla \mathcal{T}(x))^{1-\gamma} dx. \end{aligned} \quad (5.25)$$

Since the matrix $\nabla \mathcal{T}$ may not be symmetric, the functional $\mathcal{T} \mapsto \mathcal{U}[\mathcal{T} \# \varrho, \mathcal{T} \# \sigma]$ is not convex if $d \geq 2$. In order to obtain a *convex* minimization problem, we modify the functional by replacing $\nabla \mathcal{T}$ by the deformation, i.e., its symmetric part.

Definition 5.9 (Internal Energy). Let density/entropy $(\varrho, \sigma) \in \mathcal{P}_2(\mathbf{R}^d) \times \mathcal{M}_+(\mathbf{R}^d)$ be given with $\varrho = r\mathcal{L}^d$, $\sigma = \varrho S$, and $\mathcal{U}[\varrho, \sigma] < \infty$. For any $\mathcal{T} \in \mathcal{C}_\varrho$ let

$$D\mathcal{T} = \nabla\mathcal{T} \mathcal{L}^d + D^s\mathcal{T}, \quad D^s\mathcal{T} \perp \mathcal{L}^d \quad (5.26)$$

be the Lebesgue-Radon-Nikodým decomposition of its derivative. Then

$$\mathcal{U}[\mathcal{T}|_\varrho, \sigma] := \int_{\mathbf{R}^d} U(r(x), S(x)) h(\nabla\mathcal{T}(x)) dx \quad \text{for } \mathcal{T} \in \mathcal{C}_\varrho. \quad (5.27)$$

Recall that $h(\nabla\mathcal{T})$ only depends on the symmetric part of $\nabla\mathcal{T}$; see (5.7).

Remark 5.10. We have $U(r, S) \in \mathcal{L}^1(\mathbf{R}^d)$ as $\mathcal{U}[\varrho, \sigma] < \infty$. In (5.27) we may restrict the integration to $\Omega := \text{int } \overline{\text{conv}} \text{ spt } \varrho$ because the measures $\nu := U(r, S)\mathcal{L}^d$ and ϱ are mutually absolutely continuous, and $\varrho(\mathbf{R}^d \setminus \Omega) = 0$ if $\varrho \ll \mathcal{L}^d$.

Remark 5.11. Using only the symmetric part of $\nabla\mathcal{T}$ can be justified by the expectation that the map \mathcal{T} will be a perturbation of the identity, whose derivative is the identity matrix everywhere, which is symmetric. Using only $\nabla\mathcal{T}$ instead of the derivative $D\mathcal{T}$ means that the formation of vacuum does not cost any energy.

The following lemma allows us to control (5.25) in terms of (5.27).

Lemma 5.12. Suppose that density/entropy $(\varrho, \sigma) \in \mathcal{P}_2(\mathbf{R}^d) \times \mathcal{M}_+(\mathbf{R}^d)$ are given with $\varrho = r\mathcal{L}^d$, $\sigma = \varrho S$, and $\mathcal{U}[\varrho, \sigma] < \infty$. For any $\mathcal{T} \in \mathcal{C}_\varrho$ with $\mathcal{U}[\mathcal{T}|_\varrho, \sigma] < \infty$ there exists a Borel set $\Sigma \subset \mathbf{R}^d$ with $\varrho(\Sigma) = 0$ and $\mathcal{T}|_{\mathbf{R}^d \setminus \Sigma}$ injective. Then

$$\mathcal{U}[\mathcal{T} \# \varrho, \mathcal{T} \# \sigma] \leq \mathcal{U}[\mathcal{T}|_\varrho, \sigma]. \quad (5.28)$$

Proof. We have $\varrho(\mathbf{R}^d \setminus \Omega) = 0$ with $\Omega := \text{int } \overline{\text{conv}} \text{ spt } \varrho$. Since \mathcal{T} is monotone in Ω , it is differentiable for a.e. $x \in \Omega$: there exists a $(d \times d)$ -matrix $A(x)$ with

$$\lim_{x' \rightarrow x} \frac{\mathcal{T}(x') - \mathcal{T}(x) - A(x) \cdot (x' - x)}{|x' - x|} = 0; \quad (5.29)$$

see Theorem 3.2 in [1]. It follows that the function \mathcal{T} is approximately differentiable a.e. in Ω (see Definition 3.70 in [2]) and A coincides with the absolutely continuous part $\nabla\mathcal{T}$ of the derivative $D\mathcal{T}$; see Theorem 3.83 in [2] and (5.26).

Let D be the set of $x \in \Omega$ for which $\mathcal{T}(x)$ is differentiable in the sense of (5.29). Then $\mathcal{L}^d(\Omega \setminus D) = 0$. Pick a maximal monotone set-valued map u induced by a maximal monotone extension of $\Gamma := \{(x, \mathcal{T}(x)) : x \in \Omega\}$. We define

$$N := \left\{ x \in D : \text{there exists } x' \in \Omega, x' \neq x, \text{ with } \mathcal{T}(x) = \mathcal{T}(x') \right\}.$$

For $x \in N$ consider $x' \in \Omega$, $x' \neq x$, such that $\mathcal{T}(x) = \mathcal{T}(x')$. By choice of u , we get $x, x' \in u^{-1}(y)$ with $y := \mathcal{T}(x)$. Since the inverse map u^{-1} is also maximal monotone, the set $\mathcal{T}^{-1}(y)$ is closed and convex, containing with x and x' also the segment connecting the two points. Since \mathcal{T} is differentiable at x , we obtain

$$0 = \lim_{t \rightarrow 0} \frac{\mathcal{T}(x_t) - \mathcal{T}(x) - \nabla\mathcal{T}(x) \cdot (x_t - x)}{|x_t - x|} = -\nabla\mathcal{T}(x) \cdot \xi,$$

where $x_t := (1-t)x + tx'$ for $t \in [0, 1]$ and $\xi := (x' - x)/|x' - x|$. Indeed notice that $\mathcal{T}(x_t) = y$ for all $t \in [0, 1]$. Hence $\xi \neq 0$ is an eigenvector of the $(d \times d)$ -matrix $\nabla\mathcal{T}(x)$, to the eigenvalue zero. Since $x \in N$ was arbitrary, we obtain

$$N \subset \left\{ x \in D : \det \nabla\mathcal{T}(x) = 0 \right\} =: M.$$

Let $\nu := U(r, S)\mathcal{L}^d$. Since $\nu \ll \varrho$, we have that $\nu(\mathbf{R}^d \setminus \Omega) = 0$. Since $\mathcal{U}[\varrho, \sigma] < \infty$ implies that $U(r, S) \in \mathcal{L}^1(\mathbf{R}^d)$, we obtain $\nu(\Omega \setminus D) = 0$. Finally, the assumption $\mathcal{U}[\mathcal{T}|\varrho, \sigma] < \infty$ requires that $\nu(M) = 0$. We conclude that the set

$$\Sigma := (\mathbf{R}^d \setminus \Omega) \cup (\Omega \setminus D) \cup M$$

is ν -negligible, thus $\varrho(\Sigma) = 0$; see Remark 5.10. Then $\mathcal{T}|(\mathbf{R}^d \setminus \Sigma)$ is injective, which implies in particular that $\mathcal{T}\#\sigma = (S \circ \mathcal{T}^{-1})\mathcal{T}\#\varrho$. Applying Lemma 5.5.3 in [4] we conclude that the equality (5.25) is true for \mathcal{T} (with suitable modifications on sets of measure zero). We now use Lemma 5.6 to obtain the estimate

$$0 < \det \left(\epsilon(\mathcal{T}(x)) \right) \leq \det \left(\nabla \mathcal{T}(x) \right) \quad \text{for } \nu\text{-a.e. } x \in \mathbf{R}^d$$

(see also (4.16)). Then inequality (5.28) follows from the definition (5.27). \square

Proposition 5.13 (Existence of Minimizers). *Suppose that (ϱ, μ, σ) are given, with density $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$, velocity distribution $\mu \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$, and entropy $\sigma \in \mathcal{M}_+(\mathbf{R}^d)$. Assume that $\varrho = r\mathcal{L}^d$, $\sigma = \varrho S$, and $\mathcal{U}[\varrho, \sigma] < \infty$. Given any timestep $\tau > 0$, there exists a unique $\mathcal{T}_\tau \in \mathcal{C}_\varrho$ that minimizes the functional*

$$\Psi_\tau[\mathcal{T}|\mu, \sigma] := \frac{3}{4\tau^2} \int_{\mathbf{R}^{2d}} |(x + \tau\xi) - \mathcal{T}(x)|^2 \mu(dx, d\xi) + \mathcal{U}[\mathcal{T}|\varrho, \sigma] \quad (5.30)$$

for $\mathcal{T} \in \mathcal{C}_\varrho$. The minimum is finite, which implies in particular that $\mathcal{U}[\mathcal{T}_\tau|\varrho, \sigma] < \infty$. For all Borel maps $\mathcal{V}: \mathbf{R}^d \rightarrow \mathbf{R}^d$ with the property that $\mathcal{T}_\tau + \varepsilon\mathcal{V} \in \mathcal{C}_\varrho$ for some $\varepsilon > 0$, we have the following inequality: let $P(r, S) := U'(r, S)r - U(r, S)$ for $r, S \geq 0$ (where $'$ denotes differentiation with respect to r). Then

$$\begin{aligned} & -\frac{3}{2\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathcal{T}_\tau(x), \mathcal{V}(x) \rangle \mu(dx, d\xi) \\ & - \int_{\mathbf{R}^d} P(r(x), S(x)) \det \left(\epsilon(\mathcal{T}_\tau(x)) \right)^{-\gamma} \text{tr} \left(\text{cof} \left(\epsilon(\mathcal{T}_\tau(x)) \right)^T \nabla \mathcal{V}(x) \right) dx \geq 0. \end{aligned} \quad (5.31)$$

In particular, inequality (5.31) is true for $\mathcal{V} \in \mathcal{C}_\varrho$ since \mathcal{C}_ϱ is a convex cone.

Proof. We proceed in three steps.

Step 1. We observe first that the infimum $\beta := \inf_{\mathcal{T} \in \mathcal{C}_\varrho} \Psi_\tau[\mathcal{T}|\mu, \sigma]$ is nonnegative. Furthermore β is finite because we may choose $\mathcal{T} = \text{id} \in \mathcal{C}_\varrho$ to obtain

$$0 \leq \beta \leq \frac{3}{4} \int_{\mathbf{R}^{2d}} |\xi|^2 \mu(dx, d\xi) + \mathcal{U}[\varrho, \sigma] < \infty.$$

We consider a sequence of $\mathcal{T}^k \in \mathcal{C}_\varrho$ such that $\Psi_\tau[\mathcal{T}^k|\mu, \sigma] \rightarrow \beta$ as $k \rightarrow \infty$. Without loss of generality, we may assume that $\Psi_\tau[\mathcal{T}^k|\mu, \sigma] \leq \beta + 1$ for all $k \in \mathbf{N}$. Then

$$\begin{aligned} & \int_{\mathbf{R}^d} |\mathcal{T}^k(x)|^2 \varrho(dx) \\ & \leq 2 \int_{\mathbf{R}^{2d}} |(x + \tau\xi) - \mathcal{T}^k(x)|^2 \mu(dx, d\xi) + 2 \int_{\mathbf{R}^{2d}} |x + \tau\xi|^2 \mu(dx, d\xi) \\ & \leq \frac{8\tau^2}{3}(\beta + 1) + 4 \left\{ \int_{\mathbf{R}^d} |x|^2 \varrho(dx) + \tau^2 \int_{\mathbf{R}^{2d}} |\xi|^2 \mu(dx, d\xi) \right\} < \infty. \end{aligned}$$

Therefore the sequence $\{\mathcal{T}^k\}_k$ is precompact with respect to weak convergence in $\mathcal{L}^2(\mathbf{R}^d, \varrho)$: there exists a subsequence (still denoted by $\{\mathcal{T}^k\}_k$) and $\mathcal{T} \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$

such that $\mathcal{T}^k \rightharpoonup \mathcal{T}$ weakly. By Mazur's lemma, there exists a map $K: \mathbf{N} \rightarrow \mathbf{N}$ with $K(n) \geq n$ for all $n \in \mathbf{N}$, and a sequence of nonnegative numbers

$$\{\lambda_k^n: k = n \dots K(n)\}$$

with $\sum_{k=n}^{K(n)} \lambda_k^n = 1$, with the property that

$$\mathcal{S}^n := \sum_{k=n}^{K(n)} \lambda_k^n \mathcal{T}^k \rightarrow \mathcal{T} \quad \text{strongly in } \mathcal{L}^2(\mathbf{R}^d, \varrho)$$

as $n \rightarrow \infty$. Notice that $\mathcal{S}^n \in \mathcal{C}_\varrho$ since \mathcal{C}_ϱ is a convex cone. We apply Proposition 5.4 (the convexity of the quadratic term in (5.30) is easy to check) to estimate

$$\beta \leq \Psi_\tau[\mathcal{S}^n | \boldsymbol{\mu}, \sigma] \leq \sum_{k=n}^{K(n)} \lambda_k^n \Psi_\tau[\mathcal{T}^k | \boldsymbol{\mu}, \sigma] \rightarrow \beta.$$

Consequently, we obtain a strongly convergent minimizing sequence. Without loss of generality, we may assume that $\Psi_\tau[\mathcal{S}^n | \boldsymbol{\mu}, \sigma] \leq \beta + 1$ for all $n \in \mathbf{N}$. Extracting another subsequence if necessary, we may even assume the existence of a Borel set $N \subset \mathbf{R}^d$ with $\varrho(\mathbf{R}^d \setminus N) = 0$ such that $\mathcal{S}^n(x) \rightarrow \mathcal{T}(x)$ for all $x \in \mathbf{R}^d \setminus N$.

Step 2. It remains to establish the lower semicontinuity of the functional (5.30). The quadratic part is clearly lower semicontinuous with respect to weak convergence in $\mathcal{L}^2(\mathbf{R}^d, \varrho)$. For the internal energy part, we will prove that the sequence $\{\mathcal{S}^n\}_n$ is weak* precompact in $\text{BV}_{\text{loc}}(\Omega; \mathbf{R}^d)$. Then we apply Proposition 5.4.

For all $m \in \mathbf{N}$, we define the convex compact sets

$$\Omega_m := \left\{ x \in \mathbf{R}^d: |x| \leq m \text{ and } \text{dist}(x, \mathbf{R}^d \setminus \Omega) \geq 1/m \right\}.$$

Then $\bigcup_{m \in \mathbf{N}} \Omega_m = \Omega$. Let us fix m for the moment. For each $x \in \Omega_{m+1}$ there exist finitely many points in Ω with the property that x is in the interior of the convex hull of these points. Therefore we can even find an open ball centered at x that is contained in the convex hull of these points. The collection of balls obtained in this way form an open covering of Ω_{m+1} . By compactness of Ω_{m+1} , we may choose a finite subcovering. This proves the following statement: there exist finitely many points $x_m^i \in \Omega$, $i = 1 \dots I_m$ for some $I_m \in \mathbf{N}$, with the property that

$$\Omega_{m+1} \subset \text{conv } X_m, \quad \text{where } X_m := \{x_m^i: i = 1 \dots I_m\}.$$

By adapting the argument in the proof of Lemma 3.4, we can write each $x_m^i \in X_m$ as a convex combination of points $z_m^{i,j} \in \Omega \setminus N$ with $j = 1 \dots J_m^i$ for some $J_m^i \in \mathbf{N}$. Recall that $\varrho(\mathbf{R}^d \setminus N) = 0$ and $\mathcal{S}^n(x) \rightarrow \mathcal{T}(x)$ for all $x \in \mathbf{R}^d \setminus N$. Thus

$$\Omega_{m+1} \subset \text{conv } Z_m, \quad \text{where } Z_m := \{z_m^{i,j}: j = 1 \dots J_m^i, i = 1 \dots I_m\}. \quad (5.32)$$

Since the sequence $\{\mathcal{S}^n(z_m^{i,j})\}_n$ converges, it must be bounded. Let

$$C_m^n := \max_{i=1 \dots I_m} \max_{j=1 \dots J_m^i} |\mathcal{S}^n(z_m^{i,j})|.$$

Then $\{C_m^n\}_n$ is uniformly bounded for every $m \in \mathbf{N}$. We now observe that

$$\sup_{x \in \Omega_m} |\mathcal{S}^n(x)| \leq \frac{C_m^n \text{diam}(Z_m)}{\text{dist}(\Omega_m, \mathbf{R}^d \setminus \Omega_{m+1})},$$

which is bounded uniformly in n ; see Proposition 1.2 in [1] and (5.32). We conclude that $\{\mathcal{S}^n\}_n$ is uniformly bounded in $\mathcal{L}^\infty(\Omega_m; \mathbf{R}^d)$ for all $m \in \mathbf{N}$. Since

$$\int_{\Omega_m} |D\mathcal{S}^n| \leq c_d \text{diam}(\Omega_m)^{d-1} \text{osc}(\mathcal{S}^n, \Omega_m), \quad (5.33)$$

where $c_d > 0$ is a constant depending only on the space dimension, and where

$$\text{osc}(\mathcal{S}^n, A) := \sup_{x_1, x_2 \in A} |\mathcal{S}^n(x_1) - \mathcal{S}^n(x_2)| \quad \text{for all } A \subset \mathbf{R}^d$$

denotes the oscillation of \mathcal{S}^n over A , we obtain that the sequence $\{\mathcal{S}^n\}_n$ is uniformly bounded in $\text{BV}(\Omega_m; \mathbf{R}^d)$ for all $m \in \mathbf{N}$, thus precompact in $\text{BV}_{\text{loc}}(\Omega; \mathbf{R}^d)$. We refer the reader to Proposition 5.1 and Remark 5.2 in [1] for a proof of (5.33).

Extracting another subsequence if necessary (not relabeled), we find that $\mathcal{S}^n \rightharpoonup \mathcal{S}$ weak* in $\text{BV}_{\text{loc}}(\Omega; \mathbf{R}^d)$ for a suitable function $\mathcal{S} \in \text{BV}_{\text{loc}}(\Omega; \mathbf{R}^d)$. One can check that \mathcal{S} is again a monotone map on Ω (possibly after redefining \mathcal{S} on a set of measure zero). Moreover, we have $\mathcal{S}(x) = \mathcal{T}(x)$ for ϱ -a.e. $x \in \Omega$, by construction. We now define $\mathcal{T}_\tau(x) := \mathcal{S}(x)$ for $x \in \Omega$, and $\mathcal{T}_\tau(x) := 0$ for $x \in \mathbf{R}^d \setminus \Omega$. Then

$$\mathcal{T}_\tau \in \mathcal{C}_\varrho \quad \text{and} \quad \Psi_\tau[\mathcal{T}_\tau | \boldsymbol{\mu}, \sigma] \leq \liminf_{n \rightarrow \infty} \Psi_\tau[\mathcal{S}^n | \boldsymbol{\mu}, \sigma].$$

In particular, we get $\Psi_\tau[\mathcal{T}_\tau | \boldsymbol{\mu}, \sigma] = \beta$, thus \mathcal{T}_τ is a minimizer. Its uniqueness follows from the strict convexity of the first term in (5.30), which is quadratic in \mathcal{T} .

Step 3. Consider $\mathcal{V} \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$ such that $\mathcal{T}_\tau + \varepsilon \mathcal{V} \in \mathcal{C}_\varrho$ for $\varepsilon > 0$ small. Since $\mathcal{T}_\tau \in \mathcal{C}_\varrho$, we have that $\mathcal{V} \in \text{BV}_{\text{loc}}(\Omega; \mathbf{R}^d)$ as well; see Definition 5.8. Then

$$\Psi_\tau[\mathcal{T}_\tau + \varepsilon \mathcal{V} | \boldsymbol{\mu}, \sigma] - \Psi_\tau[\mathcal{T}_\tau | \boldsymbol{\mu}, \sigma] \geq 0.$$

We divide by $\varepsilon > 0$ and consider the limit $\varepsilon \rightarrow 0$. We obtain that

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0+} \frac{1}{\varepsilon} \left\{ \frac{3}{4\tau^2} \int_{\mathbf{R}^{2d}} |(x + \tau\xi) - (\mathcal{T}_\tau(x) + \varepsilon\mathcal{V}(x))|^2 \boldsymbol{\mu}(dx, d\xi) \right. \\ & \quad \left. - \frac{3}{4\tau^2} \int_{\mathbf{R}^{2d}} |(x + \tau\xi) - \mathcal{T}_\tau(x)|^2 \boldsymbol{\mu}(dx, d\xi) \right\} \\ & = -\frac{3}{2\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathcal{T}_\tau(x), \mathcal{V}(x) \rangle \boldsymbol{\mu}(dx, d\xi). \end{aligned}$$

Since $\mathcal{U}[\mathcal{T}_\tau | \varrho, \sigma] < \infty$, we can further write

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0+} \frac{\mathcal{U}[\mathcal{T}_\tau + \varepsilon \mathcal{V} | \varrho, \sigma] - \mathcal{U}[\mathcal{T}_\tau | \varrho, \sigma]}{\varepsilon} \\ & = \lim_{\varepsilon \rightarrow 0+} \int_{\mathbf{R}^d} U(r(x), S(x)) \frac{1}{\varepsilon} \left\{ \det \left(\varepsilon(\mathcal{T}_\tau(x)) + \varepsilon \varepsilon(\mathcal{V}(x)) \right)^{1-\gamma} \right. \\ & \quad \left. - \det \left(\varepsilon(\mathcal{T}_\tau(x)) \right)^{1-\gamma} \right\} dx. \end{aligned}$$

We can restrict the integration to Ω where $\nabla \mathcal{T}_\tau$, $\nabla \mathcal{V}$ are well-defined; see Remark 5.10. Since $A \mapsto (\det(A^{\text{sym}}))^{1-\gamma}$ is convex (see Proposition 5.4), the term in curly brackets is nondecreasing for a.e. $x \in \mathbf{R}^d$. By monotone convergence, it

follows that

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0+} \frac{\mathcal{U}[\mathcal{T}_\tau + \varepsilon \mathcal{V} | \varrho, \sigma] - \mathcal{U}[\mathcal{T}_\tau | \varrho, \sigma]}{\varepsilon} \\ &= - \int_{\mathbf{R}^d} P(r(x), S(x)) \det \left(\epsilon(\mathcal{T}_\tau(x)) \right)^{-\gamma} \operatorname{tr} \left(\operatorname{cof} \left(\epsilon(\mathcal{T}_\tau(x)) \right)^T \epsilon(\mathcal{V}(x)) \right) dx. \end{aligned} \quad (5.34)$$

We now can replace $\epsilon(\mathcal{V}(x))$ by $\nabla \mathcal{V}(x)$ since the antisymmetric part of the derivative cancels in the inner product with a symmetric matrix. \square

Remark 5.14. Since $\mathcal{U}[\mathcal{T}_\tau | \varrho, \sigma] < \infty$, we can apply Lemma 5.12 to conclude that \mathcal{T}_τ is essentially injective and $\mathcal{U}[\varrho_\tau, \sigma_\tau] < \infty$, where $(\varrho_\tau, \sigma_\tau) := \mathcal{T}_\tau \# (\varrho, \sigma)$. It follows that ϱ_τ must be absolutely continuous with respect to the Lebesgue measure and $\sigma_\tau = \varrho_\tau S_\tau$ with transported entropy $S_\tau := S \circ \mathcal{T}_\tau^{-1}$; recall Definition 1.1.

Remark 5.15. Using the velocities $\mathcal{V} = \pm \mathcal{T}_\tau$ in (5.31), we obtain

$$\begin{aligned} & \frac{3}{2\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathcal{T}_\tau(x), \mathcal{T}_\tau(x) \rangle \boldsymbol{\mu}(dx, d\xi) \\ &+ \int_{\mathbf{R}^d} P(r(x), S(x)) \det \left(\epsilon(\mathcal{T}_\tau(x)) \right)^{-\gamma} \operatorname{tr} \left(\operatorname{cof} \left(\epsilon(\mathcal{T}_\tau(x)) \right)^T \nabla \mathcal{T}_\tau(x) \right) dx = 0. \end{aligned} \quad (5.35)$$

This is the analogue of equality (4.4) from the pressureless case. As a consequence, we can rewrite (5.31) in the following form (cf. (4.5)): for all $\mathcal{S} \in \mathcal{C}_\varrho$ we have

$$\begin{aligned} & \frac{3}{2\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathcal{T}_\tau(x), \mathcal{S}(x) \rangle \boldsymbol{\mu}(dx, d\xi) \\ &+ \int_{\mathbf{R}^d} P(r(x), S(x)) \det \left(\epsilon(\mathcal{T}_\tau(x)) \right)^{-\gamma} \operatorname{tr} \left(\operatorname{cof} \left(\epsilon(\mathcal{T}_\tau(x)) \right)^T \nabla \mathcal{S}(x) \right) dx \leq 0. \end{aligned} \quad (5.36)$$

Using in (5.36) the constant maps $\mathcal{S}(x) = \pm b$ for all $x \in \mathbf{R}^d$, where $b \in \mathbf{R}^d$ is some vector, we conclude that the minimization in Proposition 5.13 again preserves the total momentum; see Remark 4.12 for more details.

Remark 5.16. In (5.35) we can replace $\nabla \mathcal{T}_\tau(x)$ by the deformation $\epsilon(\mathcal{T}_\tau(x))$ since the antisymmetric part cancels in the trace. By Cramer's rule, we obtain

$$\begin{aligned} & - \frac{3}{2\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathcal{T}_\tau(x), \mathcal{T}_\tau(x) \rangle \boldsymbol{\mu}(dx, d\xi) \\ &= d \int_{\mathbf{R}^d} P(r(x), S(x)) \det \left(\epsilon(\mathcal{T}_\tau(x)) \right)^{1-\gamma} dx = d(\gamma - 1) \mathcal{U}[\mathcal{T}_\tau | \varrho, \sigma]. \end{aligned} \quad (5.37)$$

Definition 5.17. For $(\varrho, \boldsymbol{\mu}, \sigma, \tau)$ as in Proposition 5.13, let \mathcal{T}_τ denote the unique minimizer considered there. We define $\mathbb{T}_\tau, \mathbb{W}_\tau, \mathbb{U}_\tau \in \mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu})$ as follows:

$$\mathbb{T}_\tau(x, \xi) := \mathcal{T}_\tau(x), \quad \mathbb{U}_\tau(x, \xi) := \mathbb{W}_\tau(x, \xi) := V_\tau(x, \xi, \mathcal{T}_\tau(x)) \quad (5.38)$$

for $\boldsymbol{\mu}$ -a.e. $(x, \xi) \in \mathbf{R}^{2d}$, with V_τ given by (3.6). Then

$$(\varrho_\tau, \sigma_\tau) := \mathcal{T}_\tau \# (\varrho, \sigma), \quad \boldsymbol{\mu}_\tau := (\mathbb{T}_\tau, \mathbb{U}_\tau) \# \boldsymbol{\mu}.$$

Remark 5.18. The definition of \mathbb{T}_τ in (5.38) is natural in view of Proposition 5.7. If $\boldsymbol{\mu} = (\operatorname{id}, \mathbf{u}) \# \varrho$ for some $\mathbf{u} \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$ and $\boldsymbol{\mu}_* := (\mathbb{T}_\tau, \mathbb{W}_\tau) \# \boldsymbol{\mu}$, then

$$\int_{\mathbf{R}^{2d}} \varphi(z, \zeta) \boldsymbol{\mu}_*(dz, d\zeta) = \int_{\mathbf{R}^d} \varphi \left(\mathcal{T}_\tau(x), \frac{3}{2} \mathcal{V}_\tau(x) - \frac{1}{2} \mathbf{u}(x) \right) \varrho(dx)$$

for all $\varphi \in \mathcal{C}_b(\mathbf{R}^{2d})$, with transport velocity $\mathcal{V}_\tau := (\mathcal{T}_\tau - \text{id})/\tau$. Let

$$\mathbf{u}_\tau(z) = \left(\frac{3}{2} \mathcal{V}_\tau - \frac{1}{2} \mathbf{u} \right) (\mathcal{T}_\tau^{-1}(z)) \quad \text{for } \varrho_\tau\text{-a.e. } z \in \mathbf{R}^d. \quad (5.39)$$

This \mathbf{u}_τ is well-defined since \mathcal{T}_τ is essentially injective; see Remark 5.14. Then

$$\int_{\mathbf{R}^{2d}} \varphi(z, \zeta) \boldsymbol{\mu}_*(dz, d\zeta) = \int_{\mathbf{R}^d} \varphi(z, \mathbf{u}_\tau(z)) \varrho_\tau(dz)$$

for all $\varphi \in \mathcal{C}_b(\mathbf{R}^d)$, which shows that $\boldsymbol{\mu}_* = (\text{id}, \mathbf{u}_\tau) \# \varrho_\tau$ and $\mathbf{u}_\tau \in \mathcal{L}^2(\mathbf{R}^d, \varrho_\tau)$. We emphasize the fact that the minimization preserves the monokinetic structure of the fluid (recall that the velocity update in (5.38) is a consequence of the minimization of the acceleration). Since the tangent cone over the cone of monotone maps at ϱ_τ equals $\mathcal{L}^2(\mathbf{R}^d, \varrho_\tau)$, no additional projection is necessary (unlike in the pressureless gas case; see Step (2) in Definition 4.9). We can therefore put $\mathbb{U}_\tau = \mathbb{W}_\tau$.

Proposition 5.19 (Stress Tensor). *Suppose that $\tau > 0$ and $(\varrho, \boldsymbol{\mu}, \sigma)$ are given, with density $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$, velocity distribution $\boldsymbol{\mu} \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$, and entropy $\sigma \in \mathcal{M}_+(\mathbf{R}^d)$. Assume that $\varrho = r\mathcal{L}^d$, $\sigma = \varrho S$, and $\mathcal{U}[\varrho, \sigma] < \infty$. Consider the unique minimizer $\mathcal{T}_\tau \in \mathcal{C}_\varrho$ from Proposition 5.13. There exists $\mathbf{M}_\tau \in \mathcal{M}(\mathbf{R}^d, \mathcal{S}_+^d)$ such that*

$$\begin{aligned} \int_{\mathbf{R}^d} \langle \epsilon(u(x)), \mathbf{M}_\tau(dx) \rangle &= -\frac{3}{2\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathcal{T}_\tau(x), u(x) \rangle \boldsymbol{\mu}(dx, d\xi) \\ &\quad - \int_{\mathbf{R}^d} P(r(x), S(x)) \det \left(\epsilon(\mathcal{T}_\tau(x)) \right)^{-\gamma} \text{tr} \left(\text{cof} \left(\epsilon(\mathcal{T}_\tau(x)) \right)^T \nabla u(x) \right) dx \end{aligned} \quad (5.40)$$

for all $u \in \mathcal{C}_*^1(\mathbf{R}^d; \mathbf{R}^d)$. In particular, we have the control

$$\begin{aligned} \int_{\mathbf{R}^d} \text{tr}(\mathbf{M}_\tau(dx)) &= -\frac{3}{2\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathcal{T}_\tau(x), x \rangle \boldsymbol{\mu}(dx, d\xi) \\ &\quad - \int_{\mathbf{R}^d} P(r(x), S(x)) \det \left(\epsilon(\mathcal{T}_\tau(x)) \right)^{-\gamma} \text{tr} \left(\text{cof} \left(\epsilon(\mathcal{T}_\tau(x)) \right)^T \right) dx. \end{aligned} \quad (5.41)$$

Proof. Since every $u \in \text{MON}(\mathbf{R}^d)$ has at most linear growth, we have $u \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$. Thus $\text{MON}(\mathbf{R}^d) \subset \mathcal{C}_\varrho$ and $\mathcal{V} := u \in \text{MON}(\mathbf{R}^d)$ is admissible in (5.31). Let

$$\mathbf{H}(dx) := P(r(x), S(x)) \det \left(\epsilon(\mathcal{T}_\tau(x)) \right)^{-\gamma} \text{cof} \left(\epsilon(\mathcal{T}_\tau(x)) \right) dx.$$

The cofactor matrix $\text{cof}(\epsilon(\mathcal{T}_\tau(x)))$ is symmetric and positive definite for a.e. $x \in \Omega$ because \mathcal{T}_τ is monotone there. Consequently, its norm can be controlled by the trace. Using $\mathcal{V} = \text{id}$ (which is an element of \mathcal{C}_ϱ) in (5.31), we obtain the estimate

$$0 \leq \int_{\mathbf{R}^d} \text{tr}(\mathbf{H}(dx)) \leq -\frac{3}{2\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau\xi) - \mathcal{T}_\tau(x), x \rangle \boldsymbol{\mu}(dx, d\xi),$$

which is finite. Thus $\mathbf{H} \in \mathcal{M}(\mathbf{R}^d; \mathcal{S}_+^d)$. If we define

$$\mathbf{F}(dx) := -\frac{3}{2\tau^2} \left((x + \tau\mathbf{u}(x)) - \mathcal{T}_\tau(x) \right) \varrho(dx),$$

where $\mathbf{u} := \mathbb{B}(\boldsymbol{\mu})$ denotes the barycentric projection of $\boldsymbol{\mu}$ (which is in $\mathcal{L}^2(\mathbf{R}^d, \varrho)$), then $\mathbf{F} \in \mathcal{M}(\mathbf{R}^d; \mathbf{R}^d)$ has finite first moment because $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$. We then apply Theorem 4.14 to obtain the representation (5.40)/(5.41); see also Remark 4.15. \square

Proposition 5.20 (Energy Balance). *Let $\tau > 0$ and $(\varrho, \mathbf{u}, \sigma)$ are given, with density $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$, Eulerian velocity $\mathbf{u} \in \mathcal{L}^2(\mathbf{R}^d, \varrho)$, and entropy $\sigma \in \mathcal{M}_+(\mathbf{R}^d)$. Suppose that $\varrho = r\mathcal{L}^d$, $\sigma = \varrho S$, and $\mathcal{U}[\varrho, \sigma] < \infty$. Let $\mathcal{T}_\tau \in \mathcal{C}_\varrho$ denote the unique minimizer from Proposition 5.13 (where $\boldsymbol{\mu} := (\text{id}, \mathbf{u})\# \varrho$) and $\mathbf{M}_\tau \in \mathcal{M}(\mathbf{R}^d; \mathcal{S}_+^d)$ the stress tensor in Proposition 5.19. Consider $(\varrho_\tau, \sigma_\tau)$ and $\boldsymbol{\mu}_\tau = (\text{id}, \mathbf{u}_\tau)\# \varrho_\tau$ as defined in the Remarks 5.14/5.18. Then we have the following energy inequality:*

$$\begin{aligned} & \int_{\mathbf{R}^d} \frac{1}{2} |\mathbf{u}_\tau(z)|^2 \varrho_\tau(dz) + \mathcal{U}[\varrho_\tau, \sigma_\tau] \\ & + \int_{\mathbf{R}^d} \text{tr}(\mathbf{M}_\tau(dx)) + \frac{1}{2} A_\tau(\boldsymbol{\mu}, \boldsymbol{\mu}_\tau)^2 + \int_0^\tau \int_{\mathbf{R}^d} \Delta_s(x) dx s ds \\ & \leq \int_{\mathbf{R}^d} \frac{1}{2} |\mathbf{u}(x)|^2 \varrho(dx) + \mathcal{U}[\varrho, \sigma], \end{aligned} \quad (5.42)$$

with dissipation of internal energy given for a.e. $x \in \mathbf{R}^d$ and $s \geq 0$ by

$$\begin{aligned} \Delta_s(x) &:= p(r(x), S(x)) \det(\epsilon(\mathcal{T}_s(x)))^{-\gamma-1} \left(\text{tr} \left(\text{cof}(\epsilon(\mathcal{T}_s(x)))^T \nabla \mathcal{V}_\tau(x) \right) \right)^2 \\ &+ P(r(x), S(x)) \det(\epsilon(\mathcal{T}_s(x)))^{-\gamma-1} \text{tr} \left(\left(\text{cof}(\epsilon(\mathcal{T}_s(x)))^T \nabla \mathcal{V}_\tau(x) \right)^2 \right). \end{aligned}$$

Here we defined the transport velocity $\mathcal{V}_\tau := (\mathcal{T}_\tau - \text{id})/\tau$ for the linear interpolation $\mathcal{T}_s := (1-s)\text{id} + s\mathcal{T}_\tau$. We also used $p(r, S) := P'(r, S)r - P(r, S)$ for $r, S \geq 0$ (where $'$ denotes differentiation with respect to r). The acceleration cost is given by

$$A_\tau(\boldsymbol{\mu}, \boldsymbol{\mu}_\tau)^2 = \int_{\mathbf{R}^d} \frac{3}{4\tau^2} |(x + \tau \mathbf{u}(x)) - \mathcal{T}_\tau(x)|^2 \varrho(dx).$$

Note that all terms on the left-hand side of (5.42) are nonnegative.

Proof. Let us first consider the kinetic energy. Because of (5.39), we have

$$\begin{aligned} & \int_{\mathbf{R}^d} \frac{1}{2} |\mathbf{u}_\tau(x)|^2 \varrho_\tau(dx) + \frac{3}{8\tau^2} \int_{\mathbf{R}^d} |(x + \tau \mathbf{u}(x)) - \mathcal{T}_\tau(x)|^2 \varrho(dx) \\ & = \int_{\mathbf{R}^d} \frac{1}{2} |\mathbf{u}(x)|^2 \varrho(dx) - \frac{3}{2\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau \xi) - \mathcal{T}_\tau(x), \mathcal{T}_\tau(x) - x \rangle \boldsymbol{\mu}(dx, d\xi) \end{aligned} \quad (5.43)$$

Combining (5.35) with the representation (5.41), we find that

$$\begin{aligned} & - \frac{3}{2\tau^2} \int_{\mathbf{R}^{2d}} \langle (x + \tau \xi) - \mathcal{T}_\tau(x), \mathcal{T}_\tau(x) - x \rangle \boldsymbol{\mu}(dx, d\xi) \\ & = \tau \int_{\mathbf{R}^d} P(r(x), S(x)) \det(\epsilon(\mathcal{T}_\tau(x)))^{-\gamma} \text{tr} \left(\text{cof}(\epsilon(\mathcal{T}_\tau(x)))^T \nabla \mathcal{V}_\tau(x) \right) dx \\ & \quad - \int_{\mathbf{R}^d} \text{tr}(\mathbf{M}_\tau(dx)). \end{aligned} \quad (5.44)$$

Taylor expansion of $s \mapsto \det(\epsilon(\mathcal{T}_s(x)))^{1-\gamma}$ around $s = \tau$ gives

$$\begin{aligned} \det\left(\epsilon(\mathcal{T}_\tau(x))\right)^{1-\gamma} &= 1 \\ &- \tau(\gamma-1) \det\left(\epsilon(\mathcal{T}_\tau(x))\right)^{-\gamma} \operatorname{tr}\left(\operatorname{cof}\left(\epsilon(\mathcal{T}_\tau(x))\right)^{\mathrm{T}} \nabla \mathcal{V}_\tau(x)\right) \\ &- \int_0^\tau \det\left(\epsilon(\mathcal{T}_s(x))\right)^{-\gamma-1} \left\{ \gamma(\gamma-1) \left(\operatorname{tr}\left(\operatorname{cof}\left(\epsilon(\mathcal{T}_s(x))\right)^{\mathrm{T}} \nabla \mathcal{V}_\tau(x)\right) \right)^2 \right. \\ &\quad \left. + (\gamma-1) \operatorname{tr}\left(\left(\operatorname{cof}\left(\epsilon(\mathcal{T}_s(x))\right)^{\mathrm{T}} \nabla \mathcal{V}_\tau(x)\right)^2\right) \right\} s \, ds \end{aligned}$$

for a.e. $x \in \Omega$. Notice that $\nabla \mathcal{T}_0(x) = \mathbb{1}$ for all $x \in \mathbf{R}^d$. We multiply by $U(r(x), S(x))$ and integrate over \mathbf{R}^d . Recalling the definition of p and P , we obtain

$$\begin{aligned} \mathcal{U}[\mathcal{T}_\tau | \varrho, \sigma] &= \mathcal{U}[\varrho, \sigma] \\ &- \tau \int_{\mathbf{R}^d} P(r(x), S(x)) \det\left(\epsilon(\mathcal{T}_\tau(x))\right)^{-\gamma} \operatorname{tr}\left(\operatorname{cof}\left(\epsilon(\mathcal{T}_\tau(x))\right)^{\mathrm{T}} \nabla \mathcal{V}_\tau(x)\right) dx \\ &- \int_0^\tau \int_{\mathbf{R}^d} \Delta_s(x) \, dx \, s \, ds. \end{aligned}$$

The first integral coincides with the first integral on the right-hand side of (5.44). Since $\mathcal{U}[\varrho_\tau, \sigma_\tau] \leq \mathcal{U}[\mathcal{T}_\tau | \varrho, \sigma]$ (see again Lemma 5.12), the result follows. \square

6. MEASURE-VALUED SOLUTIONS

In this section, we use the minimizations in Sections 4.2/5.3 to define approximate solutions to the compressible gas dynamics equations (1.1), for suitable initial data and timestep $\tau > 0$. We establish uniform bounds and prove that a subsequence converges to a measure-valued solution of (1.1) in the limit $\tau \rightarrow 0$. We will cover the pressureless case and the Euler case simultaneously, with the understanding that for the pressureless case the internal energy is set to zero. Similarly, the specific entropy is considered constant in all cases other than the full Euler case.

6.1. Interpolation. Let density $\varrho \in \mathcal{P}_2(\mathbf{R}^d)$, velocity distribution $\mu \in \mathcal{P}_\varrho(\mathbf{R}^{2d})$, and entropy $\sigma \in \mathcal{M}_+(\mathbf{R}^d)$ be given with $\mathcal{U}[\varrho, \sigma] < \infty$. The monotone transport map $\mathbb{T}_\tau \in \mathcal{L}^2(\mathbf{R}^{2d}, \mu)$ in Definitions 4.9/5.17, with $\tau > 0$, determines the new location of fluid elements at the end of the time interval $[0, \tau]$. We want to interpolate between initial and final position. We could use the path of minimal acceleration

$$Y_t(x, \xi) := x + t\xi - \left(\frac{t^2}{\tau} - \frac{t^3}{3\tau^2} \right) \frac{3}{2\tau} \left((x + \tau\xi) - \mathbb{T}_\tau(x, \xi) \right)$$

for μ -a.e. $(x, \xi) \in \mathbf{R}^{2d}$ and all $t \in [0, \tau]$. This would be the natural choice in view of the derivation of the minimal acceleration cost, which appeared in our minimization problem. Instead we prefer to use a simple linear interpolation

$$X_t(x, \xi) := x + \frac{t}{\tau} \left(\mathbb{T}_\tau(x, \xi) - x \right) \quad (6.1)$$

for μ -a.e. $(x, \xi) \in \mathbf{R}^{2d}$ and all $t \in [0, \tau]$. Recall that the linear interpolation allowed us to derive the (displacement) convexity of the internal energy and hence the

energy inequality in Proposition 5.20. The time derivatives of $X_{\tau,t}$ are then

$$\dot{X}_t(x, \xi) = \frac{\mathbb{T}_\tau(x, \xi) - x}{\tau}, \quad \ddot{X}_t(x, \xi) = 0.$$

The interpolations remain close to each other and coincide for $t = 0, \tau$ because

$$X_t(x, \xi) - Y_t(x, \xi) = t \left(1 - \frac{3t}{2\tau} + \frac{t^2}{2\tau^2} \right) \frac{1}{\tau} \left((x + \tau\xi) - \mathbb{T}_\tau(x, \xi) \right).$$

Since the $\mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu})$ -norm of the last term on the right-hand side can be estimated by the acceleration cost, which is controlled by the energy (see Propositions 4.17/5.20), the $\mathcal{L}^2(\mathbf{R}^{2d}, \boldsymbol{\mu})$ -distance between X_t and Y_t is of the order τ . Defining the densities $\rho_t := Y_t \# \boldsymbol{\mu}$ and $\varrho_t := X_t \# \boldsymbol{\mu}$, we obtain the inequality

$$W_2(\rho_t, \varrho_t) \leq t \left(1 - \frac{3t}{2\tau} + \frac{t^2}{2\tau^2} \right) \left(\frac{1}{\tau^2} \int_{\mathbf{R}^{2d}} |(x + \tau\xi) - \mathbb{T}_\tau(x, \xi)|^2 \boldsymbol{\mu}(dx, d\xi) \right)^{1/2}.$$

Definition 6.1 (Interpolation). Consider density, velocity distribution, entropy

$$\varrho \in \mathcal{P}_2(\mathbf{R}^d), \quad \boldsymbol{\mu} \in \mathcal{P}_\varrho(\mathbf{R}^{2d}), \quad \sigma \in \mathcal{M}_+(\mathbf{R}^d).$$

Suppose that $\mathcal{U}[\varrho, \sigma] < \infty$, hence $\varrho = r\mathcal{L}^d$ and $\sigma = \varrho S$ for suitable Borel functions r, S ; see Definition 1.1. Assume further that $\boldsymbol{\mu} =: (\text{id}, \mathbf{u}) \# \varrho$ with

$$\mathbf{u} \in \mathcal{L}^2(\mathbf{R}^d, \varrho) \quad \text{satisfying} \quad \int_{\mathbf{R}^d} \mathbf{u}(x) \varrho(dx) = 0.$$

For $\tau > 0$ let $(\mathbb{T}_\tau, \mathbb{W}_\tau, \mathbb{U}_\tau)$ be the maps introduced in Definitions 4.9/5.17. Let

$$(\mathcal{T}_\tau, \mathcal{W}_\tau, \mathcal{U}_\tau)(x) := (\mathbb{T}_\tau, \mathbb{W}_\tau, \mathbb{U}_\tau)(x, \mathbf{u}(x)) \quad \text{for } \varrho\text{-a.e. } x \in \mathbf{R}^d,$$

which are in $\mathcal{L}^2(\mathbf{R}^d, \varrho)$. The interpolation map satisfies

$$\begin{aligned} X_t(x, \mathbf{u}(x)) &= x + \frac{t}{\tau} (\mathcal{T}_\tau(x) - x) =: \mathcal{T}_t(x), \\ \dot{X}_t(x, \mathbf{u}(x)) &= \frac{\mathcal{T}_\tau(x) - x}{\tau} =: \mathcal{V}_\tau(x) \end{aligned} \tag{6.2}$$

for ϱ -a.e. $x \in \mathbf{R}^d$ and all $t \in [0, \tau]$. Then we define

$$(\varrho_t, \sigma_t) := \mathcal{T}_t \# (\varrho, \sigma), \quad \boldsymbol{\mu}_t := \begin{cases} \boldsymbol{\mu} & \text{if } t = 0, \\ (\mathcal{T}_t, \mathcal{V}_\tau) \# \varrho & \text{if } t \in (0, \tau), \\ (\mathcal{T}_\tau, \mathcal{U}_\tau) \# \varrho & \text{if } t = \tau. \end{cases}$$

Recall from the proof of Proposition 5.20 that in the cases with pressure

$$\mathcal{U}[\varrho_t, \sigma_t] \leq \mathcal{U}[\mathcal{T}_t | \varrho, \sigma] \leq (1 - \ell_\tau(t)) \mathcal{U}[\varrho, \sigma] + \ell_\tau(t) \mathcal{U}[\mathcal{T}_\tau | \varrho, \sigma],$$

which is finite for all $t \in [0, \tau]$. Here $\ell_\tau(t) := t/\tau$. It follows that

$$\varrho_t =: r_t \mathcal{L}^d, \quad \sigma_t = \varrho_t S_t \quad \text{with} \quad S_t := S \circ \mathcal{T}_t^{-1}.$$

Note that \mathcal{T}_t is strictly monotone and therefore invertible for all $t \in [0, \tau]$ because \mathcal{T}_τ is monotone; see (6.2). In the cases with pressure we also know that \mathcal{T}_τ is essentially injective, as shown in Lemma 5.12. We observe that the specific entropy S_t is simply transported along with the fluid. The velocity distribution $\boldsymbol{\mu}_t$ is monokinetic for all $t \in [0, \tau]$, which defines an Eulerian velocity $\mathbf{u}_t \in \mathcal{L}^2(\mathbf{R}^d, \varrho_t)$ by $\boldsymbol{\mu}_t =: (\text{id}, \mathbf{u}_t) \# \varrho_t$. This is clear by construction for $t = 0$. For $t = \tau$ we refer the reader to Definition 4.9

and Remark 5.18, respectively. For $t \in (0, \tau)$ the claim follows from the monokinetic structure of $\boldsymbol{\mu}$ and the invertibility of \mathcal{T}_t . Because of (4.14) we have

$$\int_{\mathbf{R}^d} |\mathbf{u}_t(x)|^2 \varrho_t(dx) \leq \frac{2}{3} \int_{\mathbf{R}^d} |\mathcal{W}_\tau(x)|^2 \varrho(dx) + \frac{1}{3} \int_{\mathbf{R}^d} |\mathbf{u}(x)|^2 \varrho(dx)$$

for $t \in (0, \tau)$. Using Propositions 4.17/5.20, we conclude that

$$\int_{\mathbf{R}^d} \frac{1}{2} |\mathbf{u}_t(x)|^2 \varrho_t(dx) \leq \mathcal{E}, \quad \mathcal{U}[\rho_t, \sigma_t] \leq \mathcal{E} \quad (6.3)$$

for all $t \in [0, \tau]$, with total energy $\mathcal{E} := \int_{\mathbf{R}^d} \frac{1}{2} |\mathbf{u}(x)|^2 \varrho(dx) + \mathcal{U}[\rho, \sigma]$. Note also that the momentum still has zero mean: for any vector $b \in \mathbf{R}^d$ we can write

$$\int_{\mathbf{R}^d} \langle b, \mathbf{u}_t(z) \rangle \varrho_t(dz) = \int_{\mathbf{R}^d} \langle b, \mathcal{V}_\tau(x) \rangle \varrho(dx) = \int_{\mathbf{R}^d} \langle b, \mathbf{u}(x) \rangle \varrho(dx) = 0 \quad (6.4)$$

for all $t \in (0, \tau)$; see Remarks 4.12/5.15. Similarly, we have

$$\begin{aligned} \int_{\mathbf{R}^d} \langle b, \mathbf{u}_\tau(z) \rangle \varrho_\tau(dz) &= \int_{\mathbf{R}^d} \langle b, \mathcal{U}_\tau(x) \rangle \varrho(dx) \\ &= \int_{\mathbf{R}^d} \langle b, \mathcal{W}_\tau(x) \rangle \varrho(dx) \\ &= \frac{3}{2} \int_{\mathbf{R}^d} \langle b, \mathcal{V}_\tau(x) \rangle \varrho(dx) - \frac{1}{2} \int_{\mathbf{R}^d} \langle b, \mathbf{u}(x) \rangle \varrho(dx) = 0; \end{aligned} \quad (6.5)$$

see (4.14). For the second equality we used that the barycentric projection preserves the momentum. Recall also that $\mathcal{U}_\tau = \mathcal{W}_\tau$ in the cases with pressure.

Remark 6.2. Since $\boldsymbol{\mu}_t = (\text{id}, \mathbf{u}_t) \# \varrho_t$ with $\mathbf{u}_t \in \mathcal{L}^2(\mathbf{R}^d, \varrho_t)$ for all $t \in [0, \tau]$, the map \mathbf{w} is an element of the tangent cone $\text{Tan}_{\mathbf{x}} \mathcal{C}_{\boldsymbol{\mu}_t}$; see Remark 4.10. We have

$$\left. \begin{aligned} \partial_t \varrho_t + \nabla \cdot (\varrho_t \mathbf{u}_t) &= 0 \\ \partial_t (\varrho_t \mathbf{u}_t) + \nabla \cdot (\varrho_t \mathbf{u}_t \otimes \mathbf{u}_t) &= 0 \end{aligned} \right\} \quad \text{in } \mathcal{D}'((0, \tau) \times \mathbf{R}^d).$$

6.2. Approximate Solutions. By applying the construction of the previous section to subsequent time steps, we obtain approximate solutions to (1.1) on any time interval $[0, T]$. Consider initial density, velocity distribution, and entropy

$$\bar{\varrho} \in \mathcal{P}_2(\mathbf{R}^d), \quad \bar{\boldsymbol{\mu}} \in \mathcal{P}_{\bar{\varrho}}(\mathbf{R}^{2d}), \quad \bar{\sigma} \in \mathcal{M}_+(\mathbf{R}^d).$$

Suppose $\mathcal{U}[\bar{\varrho}, \bar{\sigma}] < \infty$ so that $\bar{\varrho} =: \bar{r} \mathcal{L}^d$ and $\bar{\sigma} =: \bar{\varrho} \bar{S}$ for suitable Borel functions \bar{r}, \bar{S} ; see Definition 1.1. Assume further that $\bar{\boldsymbol{\mu}} =: (\text{id}, \bar{\mathbf{u}}) \# \bar{\varrho}$ with

$$\bar{\mathbf{u}} \in \mathcal{L}^2(\mathbf{R}^d, \bar{\varrho}) \quad \text{satisfying} \quad \int_{\mathbf{R}^d} \bar{\mathbf{u}}(x) \bar{\varrho}(dx) = 0. \quad (6.6)$$

Notice that since the hyperbolic conservation law (1.3) is invariant under transformations to a moving reference frame, the assumption (6.6) is not restrictive. It will be useful later. For later use, let us introduce the initial total energy

$$\bar{\mathcal{E}} := \int_{\mathbf{R}^d} \frac{1}{2} |\bar{\mathbf{u}}(x)|^2 \bar{\varrho}(dx) + \mathcal{U}[\bar{\varrho}, \bar{\sigma}] < \infty.$$

For given timestep $\tau > 0$ let $t_\tau^k := k\tau$ for all $k \in \mathbf{N}_0$. We define

$$\varrho_\tau^0 := \bar{\varrho}, \quad \boldsymbol{\mu}_\tau^0 := \bar{\boldsymbol{\mu}}, \quad \sigma_\tau^0 := \bar{\sigma}.$$

Then we proceed recursively: For any $k \in \mathbf{N}_0$ we define

$$\begin{aligned} \mathbb{T}_\tau^{k+1} &:= \mathbb{T}_\tau, & \mathbb{W}_\tau^{k+1} &:= \mathbb{W}_\tau, & \mathbb{U}_\tau^{k+1} &:= \mathbb{U}_\tau, \\ \varrho_\tau^{k+1} &:= \varrho_\tau, & \boldsymbol{\mu}_\tau^{k+1} &:= \boldsymbol{\mu}_\tau, & \sigma_\tau^{k+1} &:= \sigma_\tau, \end{aligned}$$

with $(\mathbb{T}_\tau, \mathbb{W}_\tau, \mathbb{U}_\tau)$ and $(\varrho_\tau, \boldsymbol{\mu}_\tau, \sigma_\tau)$ taken from Definitions 4.9/5.17, for the choice

$$\varrho := \varrho_\tau^k, \quad \boldsymbol{\mu} := \boldsymbol{\mu}_\tau^k, \quad \sigma := \sigma_\tau^k.$$

Let the maps $(r_\tau^k, \mathbf{u}_\tau^k, S_\tau^k)$ be determined by

$$\varrho =: r_\tau^k \mathcal{L}^d, \quad \boldsymbol{\mu} =: (\text{id}, \mathbf{u}_\tau^k) \# \varrho_\tau^k, \quad \sigma =: \varrho_\tau^k S_\tau^k.$$

For ϱ_τ^k -a.e. $x \in \mathbf{R}^d$ and $k \in \mathbf{N}_0$, we now define

$$(\mathcal{T}_\tau^{k+1}, \mathcal{W}_\tau^{k+1}, \mathcal{U}_\tau^{k+1})(x) := (\mathbb{T}_\tau^{k+1}, \mathbb{W}_\tau^{k+1}, \mathbb{U}_\tau^{k+1})(x, \mathbf{u}_\tau^k(x)),$$

which are in $\mathcal{L}^2(\mathbf{R}^d, \varrho_\tau^k)$. Then the interpolation is given by

$$\mathcal{T}_{\tau,t}(x) := x + \frac{t - t_\tau^k}{\tau} \mathcal{T}_\tau^{k+1}(x), \quad \mathcal{V}_\tau^{k+1}(x) := \frac{\mathcal{T}_\tau^{k+1}(x) - x}{\tau} \quad (6.7)$$

for ϱ_τ^k -a.e. $x \in \mathbf{R}^d$ and $t \in [t_\tau^k, t_\tau^{k+1})$, and we define

$$(\varrho_{\tau,t}, \sigma_{\tau,t}) := \mathcal{T}_{\tau,t} \# (\varrho_\tau^k, \sigma_\tau^k), \quad \boldsymbol{\mu}_{\tau,t} := \begin{cases} \boldsymbol{\mu}_\tau^k & \text{if } t = t_\tau^k, \\ (\mathcal{T}_{\tau,t}, \mathcal{V}_\tau^{k+1}) \# \varrho_\tau^k & \text{if } t \in (t_\tau^k, t_\tau^{k+1}). \end{cases} \quad (6.8)$$

As shown in the previous section, the internal energy $\mathcal{U}[\varrho_{\tau,t}, \sigma_{\tau,t}]$ is finite. Hence

$$\varrho_{\tau,t} =: r_{\tau,t} \mathcal{L}^d, \quad \sigma_{\tau,t} =: \varrho_{\tau,t} S_{\tau,t} \quad \text{with} \quad S_{\tau,t} = S_\tau^k \circ \mathcal{T}_{\tau,t}^{-1}.$$

Note that $\mathcal{T}_{\tau,t}$ is strictly monotone and thus invertible for all $t \geq 0$. As a consequence, the velocity distribution $\boldsymbol{\mu}_{\tau,t}$ is monokinetic, which defines

$$\mathbf{u}_{\tau,t} \in \mathcal{L}^2(\mathbf{R}^d, \varrho_{\tau,t}) \quad \text{such that} \quad \boldsymbol{\mu}_{\tau,t} =: (\text{id}, \mathbf{u}_{\tau,t}) \# \varrho_{\tau,t} \quad (6.9)$$

for $t \geq 0$. The map \mathfrak{v} is an element of the tangent cone $\text{Tan}_{\mathfrak{x}} \mathcal{C}_{\boldsymbol{\mu}_{\tau,t}}$ (see Remark 6.2), and the momentum $\mathbf{m}_{\tau,t} := \varrho_{\tau,t} \mathbf{u}_{\tau,t}$ has vanishing mean:

$$\int_{\mathbf{R}^d} \mathbf{u}_{\tau,t}(x) \varrho_{\tau,t}(dx) = 0 \quad \text{for all } t \geq 0;$$

see (6.4)/(6.5) and (6.6). Rewriting Propositions 4.17/5.20, we obtain

$$\begin{aligned} & \int_{\mathbf{R}^d} \frac{1}{2} |\mathbf{u}_\tau^{k+1}(x)|^2 \varrho_\tau^{k+1}(dx) + \mathcal{U}[\varrho_\tau^{k+1}, \sigma_\tau^{k+1}] \\ & + \int_{\mathbf{R}^d} \frac{1}{2} |\mathcal{W}_\tau^{k+1}(x) - \mathbf{u}_\tau^{k+1}(\mathcal{T}_\tau^{k+1}(x))|^2 \varrho_\tau^k(dx) \\ & + \int_{\mathbf{R}^d} \text{tr}(\mathbf{M}_\tau^{k+1}(dx)) + \int_{\mathbf{R}^d} \frac{1}{6} |\mathbf{u}_\tau^k(x) - \mathcal{W}_\tau^{k+1}(x)|^2 \varrho_\tau^k(dx) \\ & \leq \int_{\mathbf{R}^d} \frac{1}{2} |\mathbf{u}_\tau^k(x)|^2 \varrho_\tau^k(dx) + \mathcal{U}[\varrho_\tau^k, \sigma_\tau^k] \quad \text{for all } k \in \mathbf{N}_0. \end{aligned} \quad (6.10)$$

We omitted the dissipation of internal energy described in Proposition 5.20. From this and (6.3), we obtain the following energy bound: for all $t \geq 0$ we have

$$\int_{\mathbf{R}^d} \frac{1}{2} |\mathbf{u}_{\tau,t}(x)|^2 \varrho_{\tau,t}(dx) \leq \bar{\mathcal{E}}, \quad \mathcal{U}[\varrho_{\tau,t}, \sigma_{\tau,t}] \leq \bar{\mathcal{E}}. \quad (6.11)$$

We refer the reader to the discussion in the previous section for more details.

For any one-parameter family of $\varrho_t \in \mathcal{P}_2(\mathbf{R}^d)$ with $t \in [0, \infty)$ we denote by

$$\|\varrho\|_{\text{Lip}([0, \infty), \mathcal{P}_2(\mathbf{R}^d))} := \sup_{t_1 \neq t_2} \frac{W_2(\varrho_{t_1}, \varrho_{t_2})}{|t_2 - t_1|}$$

the Lipschitz seminorm, where W_2 is the Wasserstein distance.

Lemma 6.3. *For given initial data $(\bar{\varrho}, \bar{\mathbf{u}})$ as above and $\tau > 0$, consider approximate solutions $(\varrho_\tau, \mathbf{u}_\tau)$ as defined in (6.8)/(6.9). Then we have*

$$\sup_{\tau > 0} \|\varrho_\tau\|_{\text{Lip}([0, \infty), \mathcal{P}_2(\mathbf{R}^d))} \leq (2\bar{\mathcal{E}})^{1/2}.$$

The second moments remain finite: for all $\tau > 0$ and $t \in [0, \infty)$ we have

$$\left(\int_{\mathbf{R}^d} |x|^2 \varrho_{\tau, t}(dx) \right)^{1/2} \leq \left(\int_{\mathbf{R}^d} |x|^2 \bar{\varrho}(dx) \right)^{1/2} + t(2\bar{\mathcal{E}})^{1/2}.$$

Proof. We divide the proof into two steps.

Step 1. Consider first $0 \leq t_1 < t_2$ with $t_1, t_2 \in [t_\tau^k, t_\tau^{k+1})$ for some $k \in \mathbf{N}_0$. Since $\mathcal{T}_{\tau, t}(x) = x + (t - t_\tau^k) \mathcal{V}_\tau^{k+1}(x)$ for ϱ_τ^k -a.e. $x \in \mathbf{R}^d$ and $t \in [t_\tau^k, t_\tau^{k+1})$, we get

$$\begin{aligned} W_2(\varrho_{\tau, t_2}, \varrho_{\tau, t_1})^2 &\leq \int_{\mathbf{R}^d} |\mathcal{T}_{\tau, t_2}(x) - \mathcal{T}_{\tau, t_1}(x)|^2 \varrho_\tau^k(dx) \\ &= (t_2 - t_1)^2 \int_{\mathbf{R}^d} |\mathcal{V}_\tau^{k+1}(x)|^2 \varrho_\tau^k(dx). \end{aligned} \quad (6.12)$$

For every $s \in (t_\tau^k, t_\tau^{k+1})$, we have

$$\int_{\mathbf{R}^d} |\mathcal{V}_\tau^{k+1}(x)|^2 \varrho_\tau^k(dx) = \int_{\mathbf{R}^d} |\mathbf{u}_{\tau, s}(x)|^2 \varrho_{\tau, s}(dx)$$

(see (6.8)/(6.9)), which is bounded uniformly in τ, k because of (6.11). The estimate (6.12) remains true also for $t_2 = t_\tau^{k+1}$, by continuity.

Step 2. Consider now $0 \leq t_1 < t_2$ with the property that there exists at least one $k \in \mathbf{N}$ with $t_1 \leq t_\tau^k < t_2$. We use the triangle inequality to estimate

$$W_2(\varrho_{\tau, t_2}, \varrho_{\tau, t_1}) \leq W_2(\varrho_{\tau, t_2}, \varrho_\tau^{k_2}) + \sum_{k=k_1+1}^{k_2-1} W_2(\varrho_\tau^{k+1}, \varrho_\tau^k) + W_2(\varrho_\tau^{k_1+1}, \varrho_{\tau, t_1}),$$

where $k_i := \lfloor t_i/\tau \rfloor$ for $i = 1..2$. For each term, we can now apply the estimate from Step 1. Summing up all contributions, we obtain the inequality

$$W_2(\varrho_{\tau, t_2}, \varrho_{\tau, t_1}) \leq |t_2 - t_1| (2\bar{\mathcal{E}})^{1/2} \quad \text{for all } 0 \leq t_1 < t_2.$$

To control the second moments, we write

$$\begin{aligned} \left(\int_{\mathbf{R}^d} |z|^2 \varrho_{\tau, t_2}(dz) \right)^{1/2} &= \left(\int_{\mathbf{R}^{2d}} |z|^2 \gamma(dx, dz) \right)^{1/2} \\ &\leq \left(\int_{\mathbf{R}^{2d}} |x|^2 \gamma(dx, dz) \right)^{1/2} + \left(\int_{\mathbf{R}^{2d}} |z - x|^2 \gamma(dx, dz) \right)^{1/2} \\ &= \left(\int_{\mathbf{R}^d} |x|^2 \varrho_{\tau, t_1}(dx) \right)^{1/2} + W_2(\varrho_{\tau, t_2}, \varrho_{\tau, t_1}), \end{aligned}$$

with $\gamma \in \mathcal{P}_2(\mathbf{R}^{2d})$ an optimal transport plan connecting ϱ_{τ, t_1} and ϱ_{τ, t_2} . \square

6.3. Compactification. We will need compactifications of the state space.

Lemma 6.4. *Let X be a completely regular space and $\mathcal{F} \subset \mathcal{C}(X, I)$, with $I := [0, 1]$, a set of continuous functions that separates points and closed sets: for every closed set $E \subset X$ and every $x \in X \setminus E$, there exists $\Phi \in \mathcal{F}$ with $\Phi(x) \notin \overline{\Phi(E)}$. Then there exist a compact Hausdorff space \mathfrak{X} and an embedding $e: X \rightarrow \mathfrak{X}$ such that $e(X)$ is dense in \mathfrak{X} . Moreover, for any $\Phi \in \mathcal{F}$, the composition $\Phi \circ e^{-1}: e(X) \rightarrow \mathbf{R}$ has a continuous extension to all of \mathfrak{X} . If \mathcal{F} is countable, then \mathfrak{X} is metrizable.*

Proof. Consider the product space $I^{\mathcal{F}}$, which is compact in the product topology, by Tykhonov's theorem. Let the map $e: X \rightarrow I^{\mathcal{F}}$ be defined by

$$\pi_{\Phi}(e(u)) := \Phi(u) \quad \text{for all } u \in X \text{ and } \Phi \in \mathcal{F},$$

where $\pi_{\Phi}: I^{\mathcal{F}} \rightarrow I$ denotes the projection onto the Φ -component. Since \mathcal{F} separates points and closed sets, the map e is in fact an embedding (a homeomorphism between X and its image, with $e(X)$ given the relative topology of $I^{\mathcal{F}}$). We refer the reader to Proposition 4.53 of [37] for a proof. We now define \mathfrak{X} to be the closure of $e(X)$ in $I^{\mathcal{F}}$. Being a closed subset of a compact Hausdorff space, the set \mathfrak{X} is itself compact and Hausdorff. The set $e(X)$ is dense in \mathfrak{X} , by construction. We denote by \mathcal{A} the smallest closed subalgebra in $\mathcal{C}_b(X)$ containing \mathcal{F} . For any $\Phi \in \mathcal{A}$, there exists a continuous extension of $\Phi \circ e^{-1}$ to all of \mathfrak{X} ; see Proposition 4.56 in [37]. If \mathcal{F} is countable, then the set $I^{\mathcal{F}}$ is metrizable. Therefore, since every subset of a metrizable space is metrizable, we obtain that \mathfrak{X} is metrizable. We refer the reader to Section 4.8 of [37] for additional information on compactifications. \square

For simplicity of notation, we will identify X with its image $e(X)$. Then every function $\Phi \in \mathcal{A}$ can be extended as a continuous function on \mathfrak{X} . Notice that such an extension is uniquely determined because $e(X)$ is dense in \mathfrak{X} . We denote by $\mathcal{C}(\mathfrak{X})$ the space of all extensions obtained this way, and we will use the same symbols to indicate functions in \mathcal{A} and their extensions in $\mathcal{C}(\mathfrak{X})$.

6.4. Young Measures (Velocity). In this section we will introduce Young measures to capture the behavior of weakly convergent (sub)sequences of approximate momenta/velocities of (1.1) generated by our time discretization.

Let $X := \mathbf{R}^d$, which is completely regular. We introduce

$$\mathcal{W}(X) := \left\{ \Phi = \varphi \circ s : \varphi \in \mathcal{C}(\bar{\mathcal{B}}) \right\}, \quad (6.13)$$

where \mathcal{B} is the open unit ball in X and $s: X \rightarrow \mathcal{B}$ is defined by

$$s(u) := \frac{u}{\sqrt{1 + |u|^2}} \quad \text{for all } u \in X$$

(hence $s^{-1}(v) = v/\sqrt{1 - |v|^2}$ for $v \in \mathcal{B}$). One can check that $\mathcal{W}(X)$ (being isometric to $\mathcal{C}(\bar{\mathcal{B}})$) is a closed separable vector space with respect to the sup-norm. Therefore there exists a countable set \mathcal{F} that is dense in $\mathcal{W}(X) \cap \mathcal{C}(X, I)$ and separates points and closed sets. Indeed consider any closed set $E \subset X$ and any point $u \in X \setminus E$. One can find a $\Psi \in \mathcal{C}_0(X, I)$ with $\Psi(u) = 1$ and $\Psi|_E \equiv 0$, and since \mathcal{F} is dense there exists $\Phi \in \mathcal{F}$ with $\|\Phi - \Psi\|_{\mathcal{C}(X)} < \varepsilon$ for some $0 < \varepsilon < 1/2$. Applying Lemma 6.4, we obtain a compactification \mathfrak{X} (a compact, metrizable Hausdorff space) of $X := \mathbf{R}^d$. Notice that the closed subalgebra \mathcal{A} in Lemma 6.4 contains the set $\mathcal{W}(X)$.

Let now $T > 0$ be given. Then

$$\mathbb{E} := \mathcal{L}^1([0, T], \mathcal{C}_0(\mathbf{R}^d \times \mathfrak{X})) \quad (6.14)$$

is the space of (equivalence classes of) measurable maps $\phi: [0, T] \rightarrow \mathcal{C}_0(\mathbf{R}^d \times \mathfrak{X})$ (i.e., pointwise limits of sequences of simple functions) with finite norm:

$$\|\phi\|_{\mathbb{E}} := \int_0^T \|\Phi(t, \cdot)\|_{\mathcal{C}(\mathbf{R}^d \times \mathfrak{X})} dt < \infty.$$

Notice that \mathfrak{X} is compact and metrizable, hence separable. One can then show that \mathbb{E} is a separable Banach space. Its topological dual is given by

$$\mathbb{E}^* := \mathcal{L}_w^\infty([0, T], \mathcal{M}(\mathbf{R}^d \times \mathfrak{X})),$$

the space of (equivalence classes of) $\nu: [0, T] \rightarrow \mathcal{M}(\mathbf{R}^d \times \mathfrak{X})$ with

$$\begin{aligned} t \mapsto \int_{\mathbf{R}^d \times \mathfrak{X}} \phi(\mathbf{x}) \nu_t(d\mathbf{x}) \text{ measurable for all } \phi \in \mathcal{C}_0(\mathbf{R}^d \times \mathfrak{X}), \text{ and} \\ \|\nu\|_{\mathbb{E}^*} := \operatorname{ess\,sup}_{t \in [0, T]} \|\nu_t\|_{\mathcal{M}(\mathbf{R}^d \times \mathfrak{X})} < \infty \end{aligned}$$

(we write $t \mapsto \nu_t$ and $\mathbf{x} = (x, \xi)$). The duality is induced by the pairing

$$\langle \nu, \phi \rangle := \int_0^T \int_{\mathbf{R}^d \times \mathfrak{X}} \phi(t, \mathbf{x}) \nu_t(d\mathbf{x}) dt \quad (6.15)$$

for all $\phi \in \mathbb{E}$ and $\nu \in \mathbb{E}^*$. Every closed ball in \mathbb{E}^* endowed with the weak* topology is metrizable and (sequentially) compact, by Banach-Alaoglu theorem.

Recall that the functions $\mathcal{T}_{\tau, t}$ in (6.7) are well-defined ϱ_τ^k -a.e. and strictly monotone (and therefore injective) for every $t \in [t_\tau^k, t_\tau^{k+1})$ with $k \in \mathbf{N}_0$. We can therefore track the path of each fluid element starting from a generic position $\bar{x} \in \mathbf{R}^d$. By composing the transport maps of successive timesteps, we define the transport/velocity

$$X_{\tau, t} := \mathcal{T}_{\tau, t} \circ \mathcal{T}_\tau^k \circ \dots \circ \mathcal{T}_\tau^1, \quad \Xi_{\tau, t} := \begin{cases} \mathbf{u}_\tau^k \circ \mathcal{T}_\tau^k \circ \dots \circ \mathcal{T}_\tau^1 & \text{if } t = t_\tau^k, \\ \mathcal{V}_\tau^{k+1} \circ \mathcal{T}_\tau^k \circ \dots \circ \mathcal{T}_\tau^1 & \text{if } t \in (t_\tau^k, t_\tau^{k+1}), \end{cases} \quad (6.16)$$

where $k := \lfloor t/\tau \rfloor$ (the largest integer not bigger than t/τ). Notice that \mathcal{T}_τ^l is defined only ϱ_τ^l -a.e. and we may have to discard a ϱ_τ^l -null set, for every $l \in \mathbf{N}$. The preimages of these null sets under the preceding transport maps, however, can be traced back to a $\bar{\varrho}$ -negligible set, hence $X_{\tau, t}$ and $\Xi_{\tau, t}$ are well-defined $\bar{\varrho}$ -a.e. The map $t \mapsto X_{\tau, t}(\bar{x})$ is Lipschitz continuous for $\bar{\varrho}$ -a.e. $\bar{x} \in \mathbf{R}^d$ as \mathcal{V}_τ^{k+1} is finite ϱ_τ^k -a.e. We have

$$\dot{X}_{\tau, t} = \Xi_{\tau, t} \quad \text{for } t \neq t_\tau^k \text{ with } k \in \mathbf{N}_0; \quad (6.17)$$

see (6.7). To simplify the notation, let $\mathbf{X}_{\tau, t} := (X_{\tau, t}, \Xi_{\tau, t})$ for all $t \geq 0$.

For any timestep $\tau > 0$, we now define $\nu_\tau \in \mathbb{E}^*$ by

$$\int_{\mathbf{R}^d \times \mathfrak{X}} \phi(\mathbf{x}) \nu_{\tau, t}(d\mathbf{x}) := \int_{\mathbf{R}^d} \phi(x, \mathbf{u}_{\tau, t}(x)) \left(1 + |\mathbf{u}_{\tau, t}(x)|^2\right) \varrho_{\tau, t}(dx)$$

for all $\phi \in \mathcal{C}_0(\mathbf{R}^d \times \mathfrak{X})$ and $t \geq 0$. This can also be written in the form

$$\begin{aligned} & \int_{\mathbf{R}^d \times \mathfrak{X}} \phi(\mathbf{x}) \nu_{\tau,t}(d\mathbf{x}) \\ &= \begin{cases} \int_{\mathbf{R}^d} \phi(x, \mathbf{u}_\tau^k(x)) (1 + |\mathbf{u}_\tau^k(x)|^2) \varrho_\tau^k(dx) & \text{if } t = t_\tau^k, \\ \int_{\mathbf{R}^d} \phi(\mathcal{T}_{\tau,t}(x), \mathcal{V}_\tau^{k+1}(x)) (1 + |\mathcal{V}_\tau^{k+1}(x)|^2) \varrho_\tau^k(dx) & \text{if } t \in (t_\tau^k, t_\tau^{k+1}) \end{cases} \\ &= \int_{\mathbf{R}^d} \phi(\mathbf{X}_{\tau,t}(\bar{x})) (1 + |\Xi_{\tau,t}(\bar{x})|^2) \bar{\varrho}(d\bar{x}) \end{aligned} \quad (6.18)$$

for all $t \in [t_\tau^k, t_\tau^{k+1})$ and $k \in \mathbf{N}_0$, because $\boldsymbol{\mu}_{\tau,t} = \mathbf{X}_{\tau,t} \# \bar{\varrho}$ for $t \geq 0$; see (6.8). Notice that the family $\{\nu_\tau\}_{\tau>0}$ is uniformly bounded in \mathbb{E}^* : We have

$$\|\nu_{\tau,t}\|_{\mathcal{M}(\mathbf{R}^d \times \mathfrak{X})} = \int_{\mathbf{R}^d \times \mathfrak{X}} \nu_{\tau,t}(d\mathbf{x}) = 1 + \int_{\mathbf{R}^d} |\mathbf{u}_{\tau,t}(x)|^2 \varrho_{\tau,t}(dx),$$

using the fact that $\nu_{\tau,t}$ is nonnegative and (6.11). From this, we obtain the relative (sequential) compactness of $\{\nu_\tau\}_{\tau>0}$ with respect to the weak* topology.

Lemma 6.5 (Young Measure I). *Let $\tau_n \rightarrow 0$ be given and define*

$$\mathbf{X}_{n,t} := \mathbf{X}_{\tau_n,t} \quad \text{for all } t \geq 0 \text{ and } n \in \mathbf{N}.$$

Let $\nu_n \in \mathbb{E}^$ be defined as in (6.18), with $\mathbf{X}_{n,t}$ in place of $\mathbf{X}_{\tau,t}$. There exist $\nu \in \mathbb{E}^*$ and a subsequence (still denoted by $\{\nu_n\}_n$ for simplicity) such that*

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_0^T \int_{\mathbf{R}^d} \phi(t, \mathbf{X}_{n,t}(\bar{x})) (1 + |\Xi_{n,t}(\bar{x})|^2) \bar{\varrho}(d\bar{x}) dt \\ &= \int_0^T \int_{\mathbf{R}^d \times \mathfrak{X}} \phi(t, \mathbf{x}) \nu_t(d\mathbf{x}) dt \quad \text{for all } \phi \in \mathbb{E}. \end{aligned}$$

On the other hand, let us define

$$\Omega := [0, T] \times \mathbf{R}^d \quad \text{and} \quad P := \mathcal{L}^1|_{[0,T]} \otimes \bar{\varrho}. \quad (6.19)$$

We consider the flows $(t, \bar{x}) \mapsto \mathbf{X}_{\tau,t}(\bar{x})$ as elements in $\mathcal{L}^1(\Omega, P; \mathbf{R}^{2d})$. Using again (6.11) and Lemma 6.3, we conclude that the family $\{\mathbf{X}_\tau\}_{\tau>0}$ is uniformly bounded in $\mathcal{L}^2(\Omega, P; \mathbf{R}^{2d})$, hence uniformly integrable since P is a finite measure.

Definition 6.6 (Young Measures). Consider a measure space (Ω, \mathcal{S}, P) , with (Ω, P) given by (6.19) and σ -algebra \mathcal{S} . Let \mathcal{B} be the Borel σ -algebra on \mathbf{R}^{2d} . Then

$$\begin{aligned} \mathcal{Y}(\Omega, P; \mathbf{R}^{2d}) &:= \left\{ \nu \in \mathcal{M}_{+,1}(\Omega \times \mathbf{R}^{2d}, \mathcal{S} \otimes \mathcal{B}) : \right. \\ &\quad \left. \nu(A \times \mathbf{R}^{2d}) = P(A) \text{ for all } A \in \mathcal{S} \right\}, \end{aligned} \quad (6.20)$$

where $\mathcal{M}_{+,1}(\Omega \times \mathbf{R}^{2d}, \mathcal{S} \otimes \mathcal{B})$ is the space of nonnegative σ -additive measures with unit mass, defined on the product σ -algebra $\mathcal{S} \otimes \mathcal{B}$.

Lemma 6.7 (Young Measure II). *Consider (τ_n, \mathbf{X}_n) as in Lemma 6.5. There exist a Young measure $\nu \in \mathcal{Y}(\Omega, P; \mathbf{R}^{2d})$ and a subsequence (not relabeled) with*

$$\lim_{n \rightarrow \infty} \int_{\Omega} h(t, \bar{x}, \mathbf{X}_{n,t}(\bar{x})) P(dt, d\bar{x}) = \int_{\Omega} \int_{\mathbf{R}^{2d}} h(t, \bar{x}, \mathbf{x}) \nu_{(t,\bar{x})}(d\mathbf{x}) P(dt, d\bar{x})$$

for all functions (Carathéodory integrands) h on $\Omega \times \mathbf{R}^{2d}$ such that

- (1) for each $(t, \bar{x}) \in \Omega$, the map $\mathbf{x} \mapsto h(t, \bar{x}, \mathbf{x})$ is continuous on \mathbf{R}^{2d} ;
- (2) for each $\mathbf{x} \in \mathbf{R}^{2d}$, the map $(t, \bar{x}) \mapsto h(t, \bar{x}, \mathbf{x})$ is \mathcal{S} -measurable;
- (3) the sequence $\{h(\cdot, \mathbf{X}_n)\}_n$ is uniformly integrable:

$$\lim_{\alpha \rightarrow \infty} \sup_n \int_{\{|h(\cdot, \mathbf{X}_n)| \geq \alpha\}} |h(t, \bar{x}, \mathbf{X}_{n,t}(\bar{x}))| P(dt, d\bar{x}) = 0.$$

Here $\mathbf{v}(dt, d\bar{x}, d\mathbf{x}) =: v_{(t, \bar{x})}(d\mathbf{x}) P(dt, d\bar{x})$ is the disintegration of \mathbf{v} .

This is Theorem 6.3.5 in [15]. Consider now $\varphi \in \mathcal{D}([0, T], \mathbf{R}^d)$ and define

$$h(t, \bar{x}, \mathbf{x}) := g(\bar{x}) \left\{ \partial_t \varphi(t, x) + (\xi \cdot \nabla_x) \varphi(t, x) \right\}$$

for P -a.e. $(t, \bar{x}) \in \Omega$ and all $\mathbf{x} = (x, \xi) \in \mathbf{R}^{2d}$, with $g \in \mathcal{L}^\infty(\mathbf{R}^d, \bar{\varrho})$. We claim that the map $(t, \bar{x}) \mapsto h(t, \bar{x}, \mathbf{X}_{n,t}(\bar{x}))$ is uniformly integrable. Note that

$$\left\{ (t, \bar{x}) \in \Omega : |h(t, \bar{x}, \mathbf{X}_{n,t}(\bar{x}))| \geq \alpha \right\} \subset \left\{ (t, \bar{x}) \in \Omega : |\Xi_{n,t}(\bar{x})| \geq C^{-1}\alpha - 1 \right\}$$

for every $\alpha \geq 0$, where $C := \|g\|_{\mathcal{L}^\infty(\mathbf{R}^d, \bar{\varrho})} \|D\varphi\|_{\mathcal{L}^\infty([0, T] \times \mathbf{R}^d)} > 0$. (If $C = 0$, then h vanishes and there is nothing to prove.) Then we can estimate

$$\int_{\{|h(\cdot, \mathbf{X}_n)| \geq \alpha\}} |h(t, \bar{x}, \mathbf{X}_{n,t}(\bar{x}))| P(dt, d\bar{x}) \leq \frac{2C^2}{\alpha - C} \int_{\Omega} |\Xi_{n,t}(\bar{x})|^2 P(dt, d\bar{x})$$

provided $\alpha \geq 2C$. The integral on the right-hand side is bounded uniformly in n , because of (6.11). This proves the claim. On the other hand, we have

$$\begin{aligned} & \int_{\Omega} h(t, \bar{x}, \mathbf{X}_{n,t}(\bar{x})) P(dt, d\bar{x}) \\ &= \int_{\mathbf{R}^d} g(\bar{x}) \left(\int_0^T \left\{ \partial_t \varphi(t, X_{n,t}(\bar{x})) + (\Xi_{n,t}(\bar{x}) \cdot \nabla_x) \varphi(t, X_{n,t}(\bar{x})) \right\} dt \right) \bar{\varrho}(d\bar{x}) \\ &= \int_{\mathbf{R}^d} g(\bar{x}) \left(\int_0^T \frac{d}{dt} \varphi(t, X_{n,t}(\bar{x})) dt \right) \bar{\varrho}(d\bar{x}) \\ &= - \int_{\mathbf{R}^d} g(\bar{x}) \varphi(0, \bar{x}) \bar{\varrho}(d\bar{x}) \end{aligned}$$

for all n , using (6.17). From Lemma 6.7, we obtain (recall that $\mathbf{x} = (x, \xi)$)

$$\begin{aligned} & - \int_{\mathbf{R}^d} g(\bar{x}) \varphi(0, \bar{x}) \bar{\varrho}(d\bar{x}) \\ &= \int_{\mathbf{R}^d} g(\bar{x}) \left(\int_{[0, T] \times \mathbf{R}^{2d}} \left\{ \partial_t \varphi(t, x) + (\xi \cdot \nabla_x) \varphi(t, x) \right\} v_{(t, \bar{x})}(d\mathbf{x}) dt \right) \bar{\varrho}(d\bar{x}) \end{aligned}$$

for all $g \in \mathcal{L}^\infty(\mathbf{R}^d, \bar{\varrho})$ and $\varphi \in \mathcal{D}([0, T] \times \mathbf{R}^d)$. Since g is arbitrary, we get

$$-\varphi(0, \bar{x}) = \int_{[0, T] \times \mathbf{R}^{2d}} \left\{ \partial_t \varphi(t, x) + (\xi \cdot \nabla_x) \varphi(t, x) \right\} v_{(t, \bar{x})}(d\mathbf{x}) dt$$

for $\bar{\varrho}$ -a.e. $\bar{x} \in \mathbf{R}^d$. The Young measure

$$v_{(t, \bar{x})}(d\mathbf{x}) dt \quad \text{for } t \in [0, T] \text{ and } \mathbf{x} \in \mathbf{R}^{2d}$$

describes where fluid elements initially at position $\bar{x} \in \mathbf{R}^d$ travel over time. Applying the terminology of [6], we will call \mathbf{v} a transport measure. Note that our construction is also similar to the generalized flows introduced in [10]. We do not

know whether the transport measures are induced by single trajectories (as is the case for approximate solutions), or whether fluid elements may split up into a superposition of curves. The transport measure ν determines the transport of both mass:

$$-\int_{\mathbf{R}^d} \varphi(0, x) \bar{\varrho}(d\bar{x}) = \int_0^T \int_{\mathbf{R}^d} \left\{ \partial_t \varphi(t, x) \varrho_t(dx) + \nabla_x \varphi(t, x) \cdot \langle \varrho \mathbf{u} \rangle_t(dx) \right\} dt$$

$$\text{where } \begin{cases} \varrho_t := \int_{\mathbf{R}^d} \int_{\mathbf{R}^d} v_{(t, \bar{x})}(\cdot, d\xi) \bar{\varrho}(d\bar{x}), \\ \langle \varrho \mathbf{u} \rangle_t := \int_{\mathbf{R}^d} \int_{\mathbf{R}^d} \xi v_{(t, \bar{x})}(\cdot, d\xi) \bar{\varrho}(d\bar{x}), \end{cases} \quad (6.21)$$

for all $\varphi \in \mathcal{D}([0, T] \times \mathbf{R}^d)$, and entropy:

$$-\int_{\mathbf{R}^d} \varphi(0, x) \bar{\sigma}(d\bar{x}) = \int_0^T \int_{\mathbf{R}^d} \left\{ \partial_t \varphi(t, x) \sigma_t(dx) + \nabla_x \varphi(t, x) \cdot \langle \sigma \mathbf{u} \rangle_t(dx) \right\} dt$$

$$\text{where } \begin{cases} \sigma_t := \int_{\mathbf{R}^d} \bar{S}(\bar{x}) \int_{\mathbf{R}^d} v_{(t, \bar{x})}(\cdot, d\xi) \bar{\varrho}(d\bar{x}), \\ \langle \sigma \mathbf{u} \rangle_t := \int_{\mathbf{R}^d} \bar{S}(\bar{x}) \int_{\mathbf{R}^d} \xi v_{(t, \bar{x})}(\cdot, d\xi) \bar{\varrho}(d\bar{x}). \end{cases} \quad (6.22)$$

Recall that $\bar{\sigma} = \bar{\varrho} \bar{S}$, with nonnegative specific entropy $\bar{S} \in \mathcal{L}^\infty(\mathbf{R}^d, \bar{\varrho})$.

By construction, the momentum $\langle \varrho \mathbf{u} \rangle_t$ is absolutely continuous with respect to the density ϱ_t , for a.e. $t \in [0, T]$. We consider the disintegration determined by

$$\int_{\mathbf{R}^{2d}} \phi(\mathbf{x}) \int_{\mathbf{R}^d} v_{(t, \bar{x})}(d\mathbf{x}) \bar{\varrho}(d\bar{x}) =: \int_{\mathbf{R}^{2d}} \phi(x, \xi) \hat{v}_{(t, x)}(d\xi) \varrho_t(dx) \quad (6.23)$$

for all $\phi \in \mathcal{C}_b(\mathbf{R}^{2d})$ (see (6.21)), and define a velocity

$$\mathbf{u}_t(x) := \int_{\mathbf{R}^d} \xi \hat{v}_{(t, x)}(d\xi) \quad \text{for } \varrho_t\text{-a.e. } x \in \mathbf{R}^d, \quad (6.24)$$

which is the Radon-Nikodým derivative of $\langle \varrho \mathbf{u} \rangle_t$ with respect to ϱ_t . Since $\hat{v}_{(t, x)}$ is a probability measure, we can use Jensen's inequality to estimate the norm

$$\begin{aligned} \int_{\mathbf{R}^d} |\mathbf{u}_t(x)|^2 \varrho_t(dx) &\leq \int_{\mathbf{R}^d} \int_{\mathbf{R}^d} |\xi|^2 \hat{v}_{(t, x)}(d\xi) \varrho_t(dx) \\ &= \int_{\mathbf{R}^d} \int_{\mathbf{R}^{2d}} |\xi|^2 v_{(t, \bar{x})}(d\mathbf{x}, d\xi) \bar{\varrho}(d\bar{x}). \end{aligned}$$

By lower semicontinuity, the last integral is bounded by the lim inf of

$$\int_{\mathbf{R}^d} |\Xi_{n, t}(\bar{x})|^2 \bar{\varrho}(d\bar{x}) = \int_{\mathbf{R}^d} |\mathbf{u}_{\tau_n, t}(x)|^2 \varrho_{\tau_n, t}(dx)$$

(see (6.8)), which is uniformly bounded, both in n and in $t \geq 0$, because of (6.11). We conclude that $\mathbf{u}_t \in \mathcal{L}^2(\mathbf{R}^d, \varrho_t)$ for a.e. $t \in [0, T]$, and

$$\varrho \mathbf{u} \in \mathcal{L}_w^\infty([0, T], \mathcal{M}(\mathbf{R}^d; \mathbf{R}^d)).$$

Similarly, the entropy momentum $\langle \sigma \mathbf{u} \rangle_t$ is absolutely continuous with respect to σ_t , with Radon-Nikodým derivative $\hat{\mathbf{u}}_t \in \mathcal{L}^2(\mathbf{R}^d, \sigma_t)$ for a.e. $t \in [0, T]$, and

$$\sigma \hat{\mathbf{u}} \in \mathcal{L}_w^\infty([0, T], \mathcal{M}(\mathbf{R}^d; \mathbf{R}^d)).$$

Here we use again that the initial specific entropy $\bar{S} \in \mathcal{L}^\infty(\mathbf{R}^d, \bar{\varrho})$.

If the transport measure \mathbf{v} is generated by a family of nonintersecting curves, which are labeled by $\bar{x} \in \mathbf{R}^d$, then $\hat{\mathbf{u}}_t = \mathbf{u}_t$ and $\sigma_t = \varrho_t S_t$ for a.e. $t \in [0, T]$, with S_t transported along the trajectories. Otherwise, there may be mixing of entropy across intersecting paths starting from different initial positions.

The Young measures $\nu \in \mathbb{E}^*$ of Lemma 6.5 and $\mathbf{v} \in \mathcal{Y}(\Omega, P; \mathbf{R}^{2d})$ of Lemma 6.7 are related: For any $\eta \in \mathcal{C}([0, T])$ and $h \in \mathcal{C}_c(\mathbf{R}^{2d})$, the maps

$$(t, \bar{x}) \mapsto \eta(t) h(\mathbf{X}_{n,t}(\bar{x})) \quad \text{for } \bar{\varrho}\text{-a.e. } \bar{x} \in \mathbf{R}^d \text{ and } t \in [0, T]$$

are bounded pointwise, thus uniformly integrable over (Ω, P) . The map

$$\mathbf{x} \mapsto \frac{h(\mathbf{x})}{1 + |\xi|^2} \quad \text{for } \mathbf{x} = (x, \xi) \in \mathbf{R}^{2d},$$

extended by zero to the compactification \mathfrak{X} , belongs to $\mathcal{C}_0(\mathbf{R}^d \times \mathfrak{X})$. From Lemmas 6.5 and 6.7, we obtain the following equality: for all η, h as above, we have

$$\int_0^T \eta(t) \int_{\mathbf{R}^d \times \mathfrak{X}} h(\mathbf{x}) \frac{\nu_t(d\mathbf{x})}{1 + |\xi|^2} dt = \int_0^T \eta(t) \int_{\mathbf{R}^d} \int_{\mathbf{R}^{2d}} h(\mathbf{x}) v_{(t, \bar{x})}(d\mathbf{x}) \bar{\varrho}(d\bar{x}) dt.$$

Since $\eta \in \mathcal{C}([0, T])$ was arbitrary, it follows that

$$\int_{\mathbf{R}^d \times \mathfrak{X}} h(\mathbf{x}) \frac{\nu_t(d\mathbf{x})}{1 + |\xi|^2} = \int_{\mathbf{R}^d} \int_{\mathbf{R}^{2d}} h(\mathbf{x}) v_{(t, \bar{x})}(d\mathbf{x}) \bar{\varrho}(d\bar{x}) \quad (6.25)$$

for a.e. $t \in [0, T]$. We may replace $\mathbf{R}^d \times \mathfrak{X}$ by \mathbf{R}^{2d} because the integrand vanishes on $\mathbf{R}^d \times (\mathfrak{X} \setminus \mathbf{R}^d)$. Recall that both the measure ν_t and the second moment

$$\int_{\mathbf{R}^d} \int_{\mathbf{R}^{2d}} |\xi|^2 v_{(t, \bar{x})}(d\mathbf{x}) \bar{\varrho}(d\bar{x})$$

are finite. We now consider a radially symmetric and nonincreasing cut-off function $\phi \in \mathcal{C}_c(\mathbf{R}^d)$ such that $\phi(\xi) = 1$ if $|\xi| \leq 1$ and $\phi(\xi) = 0$ if $|\xi| \geq 2$. We use

$$h_r(x, \xi) := \varphi(x) \phi(r\xi) (\xi \otimes \xi) \quad \text{for } x, \xi \in \mathbf{R}^d \text{ and } r > 0$$

in (6.25) and let $r \rightarrow 0$. By dominated convergence, we obtain the equality

$$\begin{aligned} \int_{\mathbf{R}^{2d}} \varphi(x) (\xi \otimes \xi) \frac{\nu_t(d\mathbf{x})}{1 + |\xi|^2} &= \int_{\mathbf{R}^d} \int_{\mathbf{R}^{2d}} \varphi(x) (\xi \otimes \xi) v_{(t, \bar{x})}(d\mathbf{x}) \bar{\varrho}(d\bar{x}) \\ &= \int_{\mathbf{R}^d} \varphi(x) \left(\int_{\mathbf{R}^d} (\xi \otimes \xi) \hat{v}_{(t, x)}(d\xi) \right) \varrho_t(dx) \end{aligned}$$

for all $\varphi \in \mathcal{C}_c(\mathbf{R}^d)$ and a.e. $t \in [0, T]$; see (6.23). By Jensen inequality, we have

$$\mathbf{u}_t(x) \otimes \mathbf{u}_t(x) \leq \int_{\mathbf{R}^d} (\xi \otimes \xi) \hat{v}_{(t, x)}(d\xi) \quad \text{for } \varrho_t\text{-a.e. } x \in \mathbf{R}^d \quad (6.26)$$

and a.e. $t \in [0, T]$ (see (6.24)), in the sense of symmetric matrices.

We now define a map $\langle \varrho \mathbf{u} \otimes \mathbf{u} \rangle \in \mathcal{L}_w^\infty([0, T], \mathcal{M}(\mathbf{R}^d; \mathcal{S}_+^d))$ by

$$\int_{\mathbf{R}^d} \phi(x) \langle \varrho \mathbf{u} \otimes \mathbf{u} \rangle_t(dx) := \int_{\mathbf{R}^d \times \mathfrak{X}} \phi(x) (\xi \otimes \xi) \frac{\nu_t(d\mathbf{x})}{1 + |\xi|^2}$$

for all $\phi \in \mathcal{C}_c(\mathbf{R}^d)$ and a.e. $t \in [0, T]$. Using Lemma 6.5, we get

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_0^T \int_{\mathbf{R}^d} \phi(t, X_{n,t}(\bar{x})) (\Xi_{n,t}(\bar{x}) \otimes \Xi_{n,t}(\bar{x})) \bar{\varrho}(d\bar{x}) dt \\ &= \int_0^T \int_{\mathbf{R}^d} \phi(t, x) \langle \varrho \mathbf{u} \otimes \mathbf{u} \rangle_t(dx) dt \end{aligned} \quad (6.27)$$

for all $\phi \in \mathcal{C}_0([0, T] \times \mathbf{R}^d)$. We can decompose

$$\langle \varrho \mathbf{u} \otimes \mathbf{u} \rangle_t(dx) = (\mathbf{u}_t(x) \otimes \mathbf{u}_t(x)) \varrho_t(dx) + R_t(dx) \quad (6.28)$$

for a.e. $t \in [0, T]$, with rest term

$$\begin{aligned} R_t(dx) &:= \left(\int_{\mathbf{R}^d} (\xi \otimes \xi) \hat{v}_{(t,x)}(d\xi) - \mathbf{u}_t(x) \otimes \mathbf{u}_t(x) \right) \varrho_t(dx) \\ &\quad + \int_{\mathfrak{X} \setminus \mathbf{R}^d} (\xi \otimes \xi) \frac{\nu_t(dx, d\xi)}{1 + |\xi|^2}. \end{aligned} \quad (6.29)$$

Because of (6.26), we have $R \in \mathcal{L}_w^\infty([0, T], \mathcal{M}(\mathbf{R}^d; \mathcal{S}_+^d))$. In particular, the matrices are positive semidefinite. Since R captures oscillations and concentrations in the weak* convergence (6.27), we will refer to R as the Reynolds tensor field.

6.5. Young Measures (Pressure). In this section we will introduce a Young measure that captures the behavior of weakly convergent (sub)sequences of approximate pressure tensor fields of (1.1) generated by our time discretization.

Let $X := \mathcal{S}_{++}^d$ (see Section 5.1), which is completely regular. We now define

$$p(A) := \det(A)^{1-\gamma} A^{-1}, \quad h(A) := \operatorname{tr}(p(A))$$

for all $A \in X$, where $\gamma > 1$ is a constant. Notice that $h(A)$ controls the norm of $p(A)$ if A is symmetric and positive definite. Then $h(A) \rightarrow \infty$ if one of the eigenvalues of A converges to zero. For $|A| \rightarrow \infty$ we have $h(A) \rightarrow 0$. One can check that

$$\mathcal{W}(X) := \left\{ \Phi = \varphi + \frac{\alpha}{1+h} + \frac{\operatorname{tr}(pB)}{1+h} : \varphi \in \mathcal{C}_0(X), \alpha \in \mathbf{R}, B \in \mathcal{S}^d \right\} \quad (6.30)$$

is a closed separable vector space with respect to the sup-norm. Indeed the functions $1/(1+h)$ and $p/(1+h)$ span a k -dimensional vector space, $k := 1 + d(d+1)/2$, that is isomorphic to \mathbf{R}^k and has a countable dense subset, as does $\mathcal{C}_0(X)$. Therefore there exists a countable set \mathcal{F} that is dense in $\mathcal{W}(X) \cap \mathcal{C}(X, I)$ and separates points and closed sets; see Section 6.4 for additional details. Applying Lemma 6.4 again, we obtain a compactification \mathfrak{X} (a compact, metrizable Hausdorff space) of $X := \mathcal{S}_{++}^d$. Notice that the closed subalgebra \mathcal{A} in Lemma 6.4 contains the set $\mathcal{W}(X)$.

We now proceed as in Section 6.4. Let $T > 0$ be given and consider

$$\mathbb{E} := \mathcal{L}^1([0, T], \mathcal{C}_0(\mathbf{R}^d \times \mathfrak{X})),$$

with \mathfrak{X} the compactification of \mathcal{S}_{++}^d . For any $\tau > 0$, we define $\nu_\tau \in \mathbb{E}^*$ by

$$\begin{aligned} & \int_{\mathbf{R}^d \times \mathfrak{X}} \phi(x, A) \nu_{\tau,t}(dx, dA) \\ &:= \int_{\mathbf{R}^d} P(r_\tau^k(x), S_\tau^k(x)) \phi\left(x, \epsilon(\mathcal{T}_\tau^{k+1}(x))\right) \left(1 + h\left(\epsilon(\mathcal{T}_\tau^{k+1}(x))\right)\right) dx \end{aligned} \quad (6.31)$$

for all $\phi \in \mathcal{C}_0(\mathbf{R}^d \times \mathfrak{X})$ and $t \in [t_t^k, t_{\tau}^{k+1})$, $k \in \mathbf{N}_0$. Here $\varrho_{\tau}^k =: r_{\tau}^k \mathcal{L}^d$ and $\sigma_{\tau}^k =: \varrho_{\tau}^k S_{\tau}^k$. Notice that the family $\{\nu_{\tau}\}_{\tau>0}$ is uniformly bounded in \mathbb{E}^* : Let

$$\mathbf{H}_{\tau}^k(x) := P(r_{\tau}^k(x), S_{\tau}^k(x)) \det \left(\epsilon(\mathcal{T}_{\tau}^{k+1}(x)) \right)^{-\gamma} \operatorname{cof} \left(\epsilon(\mathcal{T}_{\tau}^{k+1}(x)) \right), \quad (6.32)$$

which is symmetric and positive semidefinite so that $\operatorname{tr}(\mathbf{H}_{\tau}^k(x)) \geq 0$ for a.e. $x \in \mathbf{R}^d$. We use (5.41) (with $x = \mathcal{T}_{\tau}^{k+1}(x) - \tau \mathcal{V}_{\tau}^{k+1}(x)$) and (5.37) to get

$$\begin{aligned} \int_{\mathbf{R}^d} \operatorname{tr}(\mathbf{H}_{\tau}^k(x)) \, dx &\leq d(\gamma - 1) \mathcal{U}[\mathcal{T}_{\tau}^{k+1} | \varrho_{\tau}^k, \sigma_{\tau}^k] \\ &\quad + \frac{3}{2} \left(\int_{\mathbf{R}^d} |\mathbf{u}_{\tau}^k(x) - \mathcal{V}_{\tau}^{k+1}(x)|^2 \varrho_{\tau}^k(dx) \right)^{1/2} \left(\int_{\mathbf{R}^d} |\mathcal{V}_{\tau}^{k+1}(x)|^2 \varrho_{\tau}^k(dx) \right)^{1/2}, \end{aligned} \quad (6.33)$$

which remains bounded, uniformly in τ, k ; see the proof of Proposition 5.20 and (6.10)/(6.11). On the other hand, we have the following identity:

$$\|\nu_{\tau,t}\|_{\mathcal{M}(\mathbf{R}^d \times \mathfrak{X})} = \int_{\mathbf{R}^d \times \mathfrak{X}} \nu_{\tau,t}(dx, dA) = (\gamma - 1) \mathcal{U}[\varrho_{\tau}^k, \sigma_{\tau}^k] + \int_{\mathbf{R}^d} \operatorname{tr}(\mathbf{H}_{\tau}^k(x)) \, dx,$$

using the fact that $\nu_{\tau,t}$ is nonnegative and (6.11). From this, we obtain the relative (sequential) compactness of $\{\nu_{\tau}\}_{\tau>0}$ with respect to the weak* topology.

Lemma 6.8 (Young Measure III). *Let $\tau_n \rightarrow 0$ be given and define $\nu_n := \nu_{\tau_n} \in \mathbb{E}^*$ by (6.31). There exist $\nu \in \mathbb{E}^*$ and a subsequence (still denoted by $\{\nu_n\}_n$) with*

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_0^T \int_{\mathbf{R}^d \times \mathfrak{X}} \phi(t, x, A) \nu_{n,t}(dx, dA) \, dt \\ = \int_0^T \int_{\mathbf{R}^d \times \mathfrak{X}} \phi(t, x, A) \nu_t(dx, dA) \, dt \quad \text{for all } \phi \in \mathbb{E}. \end{aligned}$$

We now define the pressure tensor field $\langle \pi \rangle \in \mathcal{L}_w^{\infty}([0, T], \mathcal{M}(\mathbf{R}^d; S_+^d))$ by

$$\int_{\mathbf{R}^d} \phi(x) \langle \pi \rangle_t(dx) := \int_{\mathbf{R}^d \times \mathfrak{X}} \phi(x) p(A) \frac{\nu_t(dx, dA)}{1 + h(A)} \quad (6.34)$$

for all $\phi \in \mathcal{C}_c(\mathbf{R}^d)$ and a.e. $t \in [0, T]$.

6.6. Global Existence. In this section, we establish the global existence of measure-valued solutions to (1.1), using the results of Sections 6.4 and 6.5.

Proof of Theorem 1.2. We divide the proof into two steps.

Step 1. Suppose that $\tau_n \rightarrow 0$ is the sequence for which the approximate solutions considered in Section 6.4 generate the Young measures in Lemmas 6.5 and 6.7. They induce the density ϱ and the entropy σ that satisfy the continuity/transport equations in (1.12); see (6.21)/(6.22). The initial data is assumed in the sense of distributions, and the velocity field \mathbf{u} in the continuity equation is given by (6.24).

To prove the momentum equation in (1.12), let us first fix some notation:

$$\begin{aligned} X_n^k &:= X_{\tau_n, t_{\tau_n}^k}, \quad (\mathcal{T}_n^{k+1}, \mathcal{V}_n^{k+1}, \mathcal{W}_n^{k+1}) := (\mathcal{T}_{\tau_n}^{k+1}, \mathcal{V}_{\tau_n}^{k+1}, \mathcal{W}_{\tau_n}^{k+1}), \\ (\varrho_n^k, \mathbf{u}_n^k, \sigma_n^k) &:= (\varrho_{\tau_n}^k, \mathbf{u}_{\tau_n}^k, \sigma_{\tau_n}^k), \quad (r_n^k, S_n^k) := (r_{\tau_n}^k, S_{\tau_n}^k) \end{aligned}$$

for $k \in \mathbf{N}_0$; see Section 6.2. With $\zeta \in \mathcal{D}([0, T] \times \mathbf{R}^d; \mathbf{R}^d)$ given, we now define

$$g_n(\bar{x}) := \int_0^T \left\{ \partial_t \zeta(t, X_{n,t}(\bar{x})) + (\Xi_{n,t}(\bar{x}) \cdot \nabla_x) \zeta(t, X_{n,t}(\bar{x})) \right\} \cdot \Xi_{n,t}(\bar{x}) \, dt \quad (6.35)$$

for $\bar{\varrho}$ -a.e. $\bar{x} \in \mathbf{R}^d$ (see also (6.17)). As shown in Section 6.4, we have that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_{\mathbf{R}^d} g_n(\bar{x}) \bar{\varrho}(d\bar{x}) \\ &= \int_0^T \int_{\mathbf{R}^d} \left\{ \partial_t \zeta(t, x) \cdot \mathbf{u}_t(x) \varrho_t(dx) + \langle \nabla_x \zeta(t, x), \langle \varrho \mathbf{u} \otimes \mathbf{u} \rangle_t(dx) \rangle \right\} dt; \end{aligned} \quad (6.36)$$

see in particular (6.21)/(6.24) and (6.27). On the other hand, we can integrate in (6.35) along the trajectory starting at $\bar{x} \in \mathbf{R}^d$ to compute the acceleration

$$g_n(\bar{x}) = \int_0^T \frac{d}{dt} \left(\zeta(t, X_{n,t}(\bar{x})) \cdot \Xi_{n,t}(\bar{x}) \right) dt.$$

Note that the map $t \mapsto \Xi_{n,t}(\bar{x})$ is piecewise constant. We obtain

$$\begin{aligned} g_n(\bar{x}) &= -\zeta(0, \bar{x}) \cdot \bar{\mathbf{u}}(\bar{x}) \\ &\quad - \sum_{k=0}^N \left(\left(\mathbf{u}_n^{k+1} \circ \mathcal{T}_n^{k+1} - \mathcal{W}_n^{k+1} \right) \cdot \zeta(t_n^{k+1}, \mathcal{T}_n^{k+1}) \right) \circ X_n^k(\bar{x}) \\ &\quad - \frac{1}{2} \sum_{k=0}^N \left(\left(\mathcal{V}_n^{k+1} - \mathbf{u}_n^k \right) \cdot \left\{ \zeta(t_n^{k+1}, \mathcal{T}_n^{k+1}) - \zeta(t_n^k, \cdot) \right\} \right) \circ X_n^k(\bar{x}) \\ &\quad - \frac{3}{2} \sum_{k=0}^N \left(\left(\mathcal{V}_n^{k+1} - \mathbf{u}_n^k \right) \cdot \zeta(t_n^k, \cdot) \right) \circ X_n^k(\bar{x}). \end{aligned} \quad (6.37)$$

Here $N := \lceil T/\tau_n \rceil - 1$ (with $\lceil \alpha \rceil$ defined as the smallest integer bigger than or equal to $\alpha \in \mathbf{R}$). While the first term on the right-hand side of (6.37) is independent of n , the remaining terms can only be controled after averaging over all trajectories. The first sum is related to the barycentric projection (see Definition 4.9) and vanishes in the cases with pressure. The second sum appears because the velocity is piecewise constant in time and therefore must be updated both at the beginning and the end of each timestep (instead of being continuously modified along the path of minimal acceleration). The third sum can be expressed in terms of the stress tensor.

We integrate (6.37) against $\bar{\varrho}$ and send $n \rightarrow \infty$. The first term on the right-hand side gives the initial data $-\int_{\mathbf{R}^d} \zeta(0, \bar{x}) \cdot \bar{\mathbf{u}}(\bar{x}) \bar{\varrho}(d\bar{x})$. We observe that

$$\begin{aligned} & - \int_{\mathbf{R}^d} \left(\left(\mathbf{u}_n^{k+1} \circ \mathcal{T}_n^{k+1} - \mathcal{W}_n^{k+1} \right) \cdot \zeta(t_n^{k+1}, \mathcal{T}_n^{k+1}) \right) \circ X_n^k(\bar{x}) \bar{\varrho}(d\bar{x}) \\ &= - \int_{\mathbf{R}^d} \left(\mathbf{u}_n^{k+1}(\mathcal{T}_n^{k+1}(x)) - \mathcal{W}_n^{k+1}(x) \right) \cdot \zeta(t_n^{k+1}, \mathcal{T}_n^{k+1}(x)) \varrho_n^k(dx) = 0 \end{aligned}$$

for any $k \in \mathbf{N}_0$; see Definition 4.9. We used that $\varrho_n^k = X_n^k \# \bar{\varrho}$. Now

$$\begin{aligned} & \left| - \frac{1}{2} \int_{\mathbf{R}^d} \left(\left(\mathcal{V}_n^{k+1} - \mathbf{u}_n^k \right) \cdot \left\{ \zeta(t_n^{k+1}, \mathcal{T}_n^{k+1}) - \zeta(t_n^k, \cdot) \right\} \right) \circ X_n^k(\bar{x}) \bar{\varrho}(d\bar{x}) \right| \\ &= \left| - \frac{1}{2} \int_{\mathbf{R}^d} \left(\mathcal{V}_n^{k+1}(x) - \mathbf{u}_n^k(x) \right) \cdot \left\{ \zeta(t_n^{k+1}, \mathcal{T}_n^{k+1}(x)) - \zeta(t_n^k, x) \right\} \varrho_n^k(dx) \right|. \end{aligned}$$

The terms in curly brackets can be estimated by

$$|\zeta(t_n^{k+1}, \mathcal{T}_n^{k+1}(x)) - \zeta(t_n^k, x)| \leq \tau_n \|D\zeta\|_{\mathcal{L}^\infty([0,T] \times \mathbf{R}^d)} \left(1 + |\mathcal{V}_n^{k+1}(x)|^2 \right)^{1/2}.$$

Since the $\mathcal{L}^2(\mathbf{R}^d, \varrho_n^k)$ -norm of \mathcal{V}_n^{k+1} is bounded uniformly in k, n (see (6.11)), there exists a constant (depending only on the initial data) such that

$$\begin{aligned} & \left| -\frac{1}{2} \int_{\mathbf{R}^d} \left(\left(\mathcal{V}_n^{k+1} - \mathbf{u}_n^k \right) \cdot \left\{ \zeta(t_n^{k+1}, \mathcal{T}_n^{k+1}) - \zeta(t_n^k, \cdot) \right\} \right) \circ X_n^k(\bar{x}) \bar{\varrho}(d\bar{x}) \right| \\ & \leq \tau_n C \|D\zeta\|_{\mathcal{L}^\infty([0, T] \times \mathbf{R}^d)} \left(\int_{\mathbf{R}^d} |\mathcal{V}_n^{k+1}(x) - \mathbf{u}_n^k(x)|^2 \varrho_n^k(dx) \right)^{1/2} \end{aligned}$$

for all $k \in \mathbf{N}_0$. Summing in k and using Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} & \left| -\frac{1}{2} \sum_{k=0}^N \int_{\mathbf{R}^d} \left(\left(\mathcal{V}_n^{k+1} - \mathbf{u}_n^k \right) \cdot \left\{ \zeta(t_n^{k+1}, \mathcal{T}_n^{k+1}) - \zeta(t_n^k, \cdot) \right\} \right) \circ X_n^k(\bar{x}) \bar{\varrho}(d\bar{x}) \right| \\ & \leq (\tau_n T)^{1/2} C \|D\zeta\|_{\mathcal{L}^\infty([0, T] \times \mathbf{R}^d)} \left(\sum_{k=0}^N \int_{\mathbf{R}^d} |\mathcal{V}_n^{k+1}(x) - \mathbf{u}_n^k(x)|^2 \varrho_n^k(dx) \right)^{1/2}, \end{aligned}$$

which converges to zero as $n \rightarrow \infty$ since the sum is bounded by the energy dissipation, uniformly in n ; see (6.10) and (4.14). Finally, we have

$$\begin{aligned} & -\frac{3}{2} \int_{\mathbf{R}^d} \left(\left(\mathcal{V}_n^{k+1} - \mathbf{u}_n^k \right) \cdot \zeta(t_n^k, \cdot) \right) \circ X_n^k(\bar{x}) \bar{\varrho}(d\bar{x}) \\ & = -\frac{3}{2} \int_{\mathbf{R}^d} \left(\mathcal{V}_n^{k+1}(x) - \mathbf{u}_n^k(x) \right) \cdot \zeta(t_n^k, x) \varrho_n^k(dx) \\ & = -\tau_n \int_{\mathbf{R}^d} \langle \epsilon(\zeta(t_n^k, x)), \mathbf{M}_n^{k+1}(dx) \rangle - \tau_n \int_{\mathbf{R}^d} \langle \nabla_x \zeta(t_n^k, x), \mathbf{H}_n^k(x) \rangle dx \end{aligned} \quad (6.38)$$

(see Proposition 5.19), with the pressure tensor field \mathbf{H}_n^k defined as in (6.32). The first term on the right-hand side of (6.38) can be estimated as

$$\left| \int_{\mathbf{R}^d} \langle \epsilon(\zeta(t_n^k, x)), \mathbf{M}_n^{k+1}(dx) \rangle \right| \leq \|\nabla_x \zeta\|_{\mathcal{L}^\infty([0, T] \times \mathbf{R}^d)} \int_{\mathbf{R}^d} \text{tr}(\mathbf{M}_n^{k+1}(dx)),$$

and the last integral is summable in k because of the energy inequality (6.10). We introduce a time integration into the pressure term in (6.38) and estimate

$$\begin{aligned} & \left| \tau_n \int_{\mathbf{R}^d} \langle \nabla_x \zeta(t_n^k, x), \mathbf{H}_n^k(x) \rangle dx - \int_{t_n^k}^{t_n^{k+1}} \int_{\mathbf{R}^d} \langle \nabla_x \zeta(t, x), \mathbf{H}_n^k(x) \rangle dx dt \right| \\ & \leq \tau_n \omega(\tau_n, \nabla_x \zeta) \int_{\mathbf{R}^d} \text{tr}(\mathbf{H}_n^k(x)) dx, \end{aligned}$$

with modulus of continuity

$$\omega(h, \varphi) := \sup_{t \in [0, T-h]} \sup_{s \in [0, h]} \|\varphi(t+s, \cdot) - \varphi(t, \cdot)\|_{\mathcal{L}^\infty(\mathbf{R}^d)}$$

for all $h > 0$ and $\varphi \in \mathcal{D}([0, T] \times \mathbf{R}^d)$. We have $\omega(h, \varphi) \rightarrow 0$ as $h \rightarrow 0$.

Therefore, if we define $\mathbf{H}_n \in \mathcal{L}_w^\infty([0, T], \mathcal{M}(\mathbf{R}^d; \mathcal{S}_+^d))$ by

$$\mathbf{H}_{n,t}(dx) := \mathbf{H}_n^k(x) dx \quad \text{for } t \in [t_n^k, t_n^{k+1}) \text{ and } k \in \mathbf{N}_0,$$

then $\{\mathbf{H}_n\}_n$ is a bounded sequence. Extracting another subsequence if necessary (still denoted by $\{\mathbf{H}_n\}_n$ for simplicity) we may assume that

$$\mathbf{H}_n \rightharpoonup \langle \pi \rangle \quad \text{weak* in } \mathcal{L}_w^\infty([0, T], \mathcal{M}(\mathbf{R}^d; \mathcal{S}_+^d))$$

(defined by testing against functions in $\mathcal{L}^1([0, T], \mathcal{C}_0(\mathbf{R}^d; \mathcal{S}_+^d))$), with the pressure tensor $\langle \boldsymbol{\pi} \rangle$ generated by the Young measure of Lemma 6.8; see (6.34). We have

$$\begin{aligned} & \left| -\frac{3}{2} \sum_{k=0}^N \left(\left(\mathcal{V}_n^{k+1} - \mathbf{u}_n^k \right) \cdot \zeta(t_n^k, \cdot) \right) \circ X_n^k(\bar{x}) \bar{\varrho}(d\bar{x}) \right. \\ & \quad \left. + \int_0^T \int_{\mathbf{R}^d} \langle \nabla_x \zeta(t, x), \mathbf{H}_{n,t}(dx) \rangle dt \right| \\ & \leq \tau_n \|\nabla \zeta\|_{\mathcal{L}^\infty([0, T] \times \mathbf{R}^d)} \sum_{k=0}^N \int_{\mathbf{R}^d} \text{tr}(\mathbf{M}_n^{k+1}(dx)) + CT\omega(\tau_n, \nabla_x \zeta), \end{aligned}$$

with C some constant that depends only on the initial data. The right-hand side converges to zero as $n \rightarrow \infty$. Collecting all terms, we obtain

$$\begin{aligned} & \int_0^T \int_{\mathbf{R}^d} \left\{ \partial_t \zeta(t, x) \cdot \mathbf{u}_t(x) \varrho_t(dx) + \langle \nabla_x \zeta(t, x), \langle \varrho \mathbf{u} \otimes \mathbf{u} \rangle_t(dx) \rangle \right\} dt \\ & = - \int_{\mathbf{R}^d} \zeta(0, \bar{x}) \cdot \bar{\mathbf{u}}(\bar{x}) \bar{\varrho}(d\bar{x}) - \int_0^T \int_{\mathbf{R}^d} \langle \nabla_x \zeta(t, x), \langle \boldsymbol{\pi} \rangle_t(dx) \rangle dt \end{aligned}$$

for all $\zeta \in \mathcal{D}([0, T] \times \mathbf{R}^d; \mathbf{R}^d)$. In the pressureless case, the pressure tensor field $\langle \boldsymbol{\pi} \rangle$ vanishes. This proves the momentum equation in (1.12).

Step 3. It only remains to show the regularity statement (1.11). For the Lipschitz continuity of the density ϱ with respect to the Wasserstein distance, we refer the reader to Theorem 8.3.1 in [4] (see also Lemma 6.3). Note that ϱ satisfies a continuity equation whose velocity field \mathbf{u} has finite kinetic energy, uniformly in time.

The same argument applies for the entropy σ , which satisfies a transport equation with a different velocity field $\hat{\mathbf{u}}$. Since the initial specific entropy \bar{S} is assumed to be bounded, the velocity $\hat{\mathbf{u}}$ has again finite energy, uniformly in time.

For the Lipschitz continuity of $\mathbf{m} = \varrho \mathbf{u}$ with respect to the Kantorovich norm, we argue as follows: Notice first that the total momentum vanishes:

$$\int_{\mathbf{R}^d} \mathbf{u}_t(x) \varrho_t(dx) = 0 \quad \text{for a.e. } t \in [0, T]$$

since the same is true for the approximate solutions. Recall that $\int_{\mathbf{R}^d} \bar{\mathbf{u}}(x) \bar{\varrho}(dx) = 0$. Moreover, the momentum has finite first moment, uniformly in time:

$$\int_{\mathbf{R}^d} (1 + |x|) |\mathbf{u}_t(x)| \varrho_t(dx) \leq \left(\int_{\mathbf{R}^d} (1 + |x|)^2 \varrho_t(dx) \right)^{1/2} \left(\int_{\mathbf{R}^d} |\mathbf{u}_t(x)|^2 \varrho_t(dx) \right)^{1/2} \quad (6.39)$$

for a.e. $x \in [0, T]$, which follows by lower semicontinuity from (6.11) and Lemma 6.3. Finally, we proved in Step 1 that the momentum satisfies the equation

$$\partial_t(\varrho \mathbf{u}) = -\nabla \cdot \hat{\mathbf{M}} \quad \text{in } \mathcal{D}'([0, T] \times \mathbf{R}^d), \quad (6.40)$$

with tensor $\hat{\mathbf{M}} \in \mathcal{L}_w^\infty([0, T], \mathcal{M}(\mathbf{R}^d; \mathcal{S}_+^d))$ given by $\hat{\mathbf{M}} := \langle \varrho \mathbf{u} \otimes \mathbf{u} \rangle + \langle \boldsymbol{\pi} \rangle$. Consider generic times $0 \leq t_1 < t_2 \leq T$, and let $\zeta: \mathbf{R}^d \rightarrow \mathbf{R}^d$ with $\|\zeta\|_{\text{Lip}(\mathbf{R}^d)} \leq 1$ be given. Notice that ζ has at most linear growth at infinity, so that the integral of ζ against $\mathbf{m}_t = \varrho_t \mathbf{u}_t$ is well-defined. On the other hand, since the pairing of $D\zeta \in \mathcal{L}^\infty(\mathbf{R}^d)$

with the Borel measure $\hat{\mathbf{M}}_t$ may not be well-defined, we need a regularization. We choose a standard mollifier φ_ε and let $\zeta_\varepsilon := \varphi_\varepsilon \star \zeta$. Then we can estimate

$$|\zeta(x) - \zeta_\varepsilon(x)| = \left| \int_{\mathbf{R}^d} (\zeta(x) - \zeta(y)) \varphi_\varepsilon(x - y) dy \right| \leq C\varepsilon \quad (6.41)$$

for $x \in \mathbf{R}^d$ and $\varepsilon > 0$, where $C > 0$ is some constant. It follows that

$$\left| \int_{\mathbf{R}^d} \langle \zeta(x), \mathbf{m}_{t_2}(dx) - \mathbf{m}_{t_1}(dx) \rangle \right| \leq \left| \int_{\mathbf{R}^d} \langle \zeta_\varepsilon(x), \mathbf{m}_{t_2}(dx) - \mathbf{m}_{t_1}(dx) \rangle \right| + 2\hat{C}\varepsilon,$$

for some constant $\hat{C} > 0$ depending on (6.41) and the bound (6.39), which is uniform in time. Using a standard truncation argument and dominated convergence in (6.40) (note that $\hat{\mathbf{M}}_t$ is a finite measure for a.e. $t \in [0, T]$), we can then estimate

$$\left| \int_{\mathbf{R}^d} \langle \zeta_\varepsilon(x), \mathbf{m}_{t_2}(dx) - \mathbf{m}_{t_1}(dx) \rangle \right| \leq |t_2 - t_1| \operatorname{ess\,sup}_{t \in [0, T]} \int_{\mathbf{R}^d} \operatorname{tr}(\hat{\mathbf{M}}_t(dx)). \quad (6.42)$$

We used that $\|\zeta\|_{\operatorname{Lip}(\mathbf{R}^d)} \leq 1$, therefore $\|\zeta_\varepsilon\|_{\operatorname{Lip}(\mathbf{R}^d)} \leq 1$ as well. Since the right-hand side of (6.42) does not depend on ε , we let $\varepsilon \rightarrow 0$ and obtain the same control with ζ_ε replaced by ζ . Taking the sup over all ζ with $\|\zeta\|_{\operatorname{Lip}(\mathbf{R}^d)} \leq 1$, we get

$$\|\mathbf{m}_{t_2} - \mathbf{m}_{t_1}\|_{\mathcal{M}_K(\mathbf{R}^d)} \leq C|t_2 - t_1| \quad \text{for a.e. } t_1, t_2 \in [0, T],$$

with $C > 0$ some constant that depends only on the initial data. The estimate can be extended to all times in $[0, T]$, by continuity (possibly redefining \mathbf{m} in time on a set of measure zero). This proves the Lipschitz continuity of the momentum. \square

REFERENCES

- [1] G. Alberti and L. Ambrosio, *A geometrical approach to monotone functions in \mathbf{R}^n* , Math. Z. **230** (1999), no. 2, 259–316.
- [2] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 2000.
- [3] L. Ambrosio and N. Gigli, *Construction of the parallel transport in the Wasserstein space*, Methods Appl. Anal. **15** (2008), no. 1, 1–29.
- [4] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Lectures in Mathematics, Birkhäuser Verlag, Basel, 2008.
- [5] H. H. Bauschke and X. Wang, *The kernel average for two convex functions and its application to the extension and representation of monotone operators*, Trans. Amer. Math. Soc. **361** (2009), no. 11.
- [6] P. Bernard, *Young measures, superposition and transport*, Indiana Univ. Math. J. **57** (2008), no. 1, 247–275.
- [7] J. M. Borwein, *Maximality of sums of two maximal monotone operators in general Banach space*, Proc. Amer. Math. Soc. **135** (2007), no. 12, 3917–3924.
- [8] F. Bouchut and F. James, *Équations de transport unidimensionnelles à coefficients discontinus*, C. R. Acad. Sci. Paris Sér. I Math. **320** (1995), no. 9, 1097–1102.
- [9] ———, *Duality solutions for pressureless gases, monotone scalar conservation laws, and uniqueness*, Comm. Partial Differential Equations **24** (1999), no. 11–12, 2173–2189.
- [10] Y. Brenier, *Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations*, Comm. Pure Appl. Math. **52** (1999), no. 4, 411–452.
- [11] ———, *L^2 formulation of multidimensional scalar conservation laws*, Arch. Ration. Mech. Anal. **193** (2009), no. 1, 1–19.
- [12] Y. Brenier, W. Gangbo, G. Savaré, and M. Westdickenberg, *Sticky particle dynamics with interactions*, J. Math. Pures Appl. (9) **99** (2013), no. 5, 577–617.
- [13] Y. Brenier and E. Grenier, *Sticky particles and scalar conservation laws*, SIAM J. Numer. Anal. **35** (1998), no. 6, 2317–2328 (electronic).

- [14] A. Bressan and T. Nguyen, *Non-existence and non-uniqueness for multidimensional sticky particle systems*, Kinet. Relat. Models **7** (2014), no. 2, 205–218.
- [15] C. Castaing, P. Raynaud de Fitte, and M. Valadier, *Young measures on topological spaces*, Mathematics and its Applications, vol. 571, Kluwer Academic Publishers, Dordrecht, 2004. With applications in control theory and probability theory.
- [16] F. Cavalletti, M. Sedjro, and M. Westdickenberg, *A Simple Proof of Global Existence for the 1D Pressureless Gas Dynamics Equations*, SIAM Math. Anal. **44** (2015), no. 1, 66–79.
- [17] F. Cavalletti and M. Westdickenberg, *The polar cone of the set of monotone maps*, Proc. Amer. Math. Soc. **143** (2015), 781–787.
- [18] G.-Q. Chen, *Compactness methods and nonlinear hyperbolic conservation laws*, Some current topics on nonlinear conservation laws, 2000, pp. 33–75.
- [19] G.-Q. Chen and P. G. LeFloch, *Compressible Euler equations with general pressure law*, Arch. Ration. Mech. Anal. **153** (2000), no. 3, 221–259.
- [20] G.-Q. Chen and M. Perepelitsa, *Vanishing viscosity limit of the Navier-Stokes equations to the Euler equations for compressible fluid flow*, Comm. Pure Appl. Math. **63** (2010), no. 11, 1469–1504.
- [21] E. Chiodaroli, *A Counterexample to Well-Posedness of Entropy Solutions to the Compressible Euler System*, Accepted for publication in J. Hyperbolic Differ. Equ. (2014).
- [22] E. Chiodaroli and O. Kreml, *On the Energy Dissipation Rate of Solutions to the Compressible Isentropic Euler System*, Accepted for publication in Arch. Ration. Mech. Anal. (2014).
- [23] E. Chiodaroli, C. De Lellis, and O. Kreml, *Global Ill-Posedness of the Isentropic System of Gas Dynamics*, Accepted for publication in Comm. Pure App. Math. (2014).
- [24] I. Chitescu, R. Miculescu, L. Nita, and L. Ioana, *Monge-Kantorovich norms on spaces of vector measures* (2014), available at [arXiv:1404.4980\[math.FA\]](https://arxiv.org/abs/1404.4980).
- [25] C. De Lellis and Jr. Székelyhidi L., *The Euler equations as a differential inclusion*, Ann. of Math. (2) **170** (2009), no. 3, 1417–1436.
- [26] ———, *On admissibility criteria for weak solutions of the Euler equations*, Arch. Ration. Mech. Anal. **195** (2010), no. 1, 225–260.
- [27] C. M. Dafermos, *The entropy rate admissibility criterion for solutions of hyperbolic conservation laws*, J. Differential Equations **14** (1973), 202–212.
- [28] S. Demoulini, D. M. A. Stuart, and A. E. Tzavaras, *A variational approximation scheme for three-dimensional elastodynamics with polyconvex energy*, Arch. Ration. Mech. Anal. **157** (2001), no. 4, 325–344.
- [29] ———, *Weak-strong uniqueness of dissipative measure-valued solutions for polyconvex elastodynamics*, Arch. Ration. Mech. Anal. **205** (2012), no. 3, 927–961.
- [30] E. De Giorgi, *New problems on minimizing movements*, Boundary value problems for partial differential equations and applications, 1993, pp. 81–98.
- [31] X. X. Ding, G.-Q. Chen, and P. Z. Luo, *Convergence of the Lax-Friedrichs scheme for the system of equations of isentropic gas dynamics. I*, Acta Math. Sci. (Chinese) **7** (1987), no. 4, 467–480.
- [32] ———, *Convergence of the Lax-Friedrichs scheme for the system of equations of isentropic gas dynamics. II*, Acta Math. Sci. (Chinese) **8** (1988), no. 1, 61–94.
- [33] R. J. DiPerna, *Convergence of the viscosity method for isentropic gas dynamics*, Comm. Math. Phys. **91** (1983), no. 1, 1–30.
- [34] W. E, Yu. G. Rykov, and Ya. G. Sinai, *Generalized variational principles, global weak solutions and behavior with random initial data for systems of conservation laws arising in adhesion particle dynamics*, Comm. Math. Phys. **177** (1996), no. 2, 349–380.
- [35] E. Feireisl, *Maximal dissipation and well-posedness for the compressible Euler system*, J. Math. Fluid Mech. **16** (2014), no. 3, 447–461.
- [36] U. S. Fjordholm, R. Kaeppli, S. Mishra, and E. Tadmor, *Construction of approximate entropy measure valued solutions for hyperbolic systems of conservation laws*, Preprint (2014).
- [37] G. B. Folland, *Real analysis*, Second, Pure and Applied Mathematics (New York), John Wiley & Sons Inc., New York, 1999. Modern techniques and their applications, A Wiley-Interscience Publication.
- [38] W. Gangbo, T. Nguyen, and A. Tudorascu, *Euler-Poisson systems as action-minimizing paths in the Wasserstein space*, Arch. Ration. Mech. Anal. **192** (2009), no. 3, 419–452.
- [39] W. Gangbo and M. Westdickenberg, *Optimal transport for the system of isentropic Euler equations*, Comm. PDE **34** (2009), no. 9, 1041–1073.

- [40] N. Ghoussoub, *A variational theory for monotone vector fields*, J. Fixed Point Theory Appl. **4** (2008), no. 1, 107–135.
- [41] N. Gigli, *On the geometry of the space of probability measures endowed with the quadratic optimal transport distance*, Ph.D. Thesis, 2004.
- [42] E. Grenier, *Existence globale pour le système des gaz sans pression*, C. R. Acad. Sci. Paris Sér. I Math. **321** (1995), no. 2, 171–174.
- [43] N. J. Higham, *Computing a nearest symmetric positive semidefinite matrix*, Linear Algebra Appl. **103** (1988), 103–118.
- [44] F. Huang and Z. Wang, *Well posedness for pressureless flow*, Comm. Math. Phys. **222** (2001), no. 1, 117–146.
- [45] R. Jordan, D. Kinderlehrer, and F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. **29** (1998), no. 1, 1–17.
- [46] J. Kristensen and F. Rindler, *Characterization of generalized gradient Young measures generated by sequences in $W^{1,1}$ and BV*, Arch. Ration. Mech. Anal. **197** (2010), no. 2, 539–598.
- [47] P. G. LeFloch and M. Westdickenberg, *Finite energy solutions to the isentropic Euler equations with geometric effects*, J. Math. Pures Appl. (9) **88** (2007), no. 5, 389–429.
- [48] H. Lim, Y. Yu, J. Glimm, X. L. Li, and D. H. Sharp, *Chaos, transport and mesh convergence for fluid mixing*, Acta Math. Appl. Sin. Engl. Ser. **24** (2008), no. 3, 355–368.
- [49] H. Lim, J. Iwerks, J. Glimm, and D. H. Sharp, *Nonideal Rayleigh-Taylor mixing*, Proc. Natl. Acad. Sci. USA **107** (2010), no. 29, 12786–12792.
- [50] P.-L. Lions, B. Perthame, and P. E. Souganidis, *Existence and stability of entropy solutions for the hyperbolic systems of isentropic gas dynamics in Eulerian and Lagrangian coordinates*, Comm. Pure Appl. Math. **49** (1996), no. 6, 599–638.
- [51] P.-L. Lions, B. Perthame, and E. Tadmor, *Kinetic formulation of the isentropic gas dynamics and p-systems*, Comm. Math. Phys. **163** (1994), no. 2, 415–431.
- [52] M. Mandelkern, *Metrization of the one-point compactification*, Proc. Amer. Math. Soc. **107** (1989), no. 4, 1111–1115.
- [53] O. Moutsinga, *Convex hulls, sticky particle dynamics and pressure-less gas system*, Ann. Math. Blaise Pascal **15** (2008), no. 1, 57–80.
- [54] L. Natile and G. Savaré, *A Wasserstein approach to the one-dimensional sticky particle system*, SIAM J. Math. Anal. **41** (2009), no. 4, 1340–1365.
- [55] T. Nguyen and A. Tudorascu, *Pressureless Euler/Euler-Poisson systems via adhesion dynamics and scalar conservation laws*, SIAM J. Math. Anal. **40** (2008), no. 2, 754–775.
- [56] F. Poupaud and M. Rasle, *Measure solutions to the linear multi-dimensional transport equation with non-smooth coefficients*, Comm. Partial Differential Equations **22** (1997), no. 1-2, 337–358.
- [57] F. Rindler, *Lower semicontinuity for integral functionals in the space of functions of bounded deformation via rigidity and Young measures*, Arch. Ration. Mech. Anal. **202** (2011), no. 1, 63–113.
- [58] R. T. Rockafellar, *On the maximality of sums of nonlinear monotone operators*, Trans. Amer. Math. Soc. **149** (1970), 75–88.
- [59] ———, *Convex analysis*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks.
- [60] M. Sever, *An existence theorem in the large for zero-pressure gas dynamics*, Differential Integral Equations **14** (2001), no. 9, 1077–1092.
- [61] E. Tadmor, *A minimum entropy principle in the gas dynamics equations*, Appl. Numer. Math. **2** (1986), no. 3-5, 211–219.
- [62] M. Westdickenberg, *Projections onto the cone of optimal transport maps and compressible fluid flows*, J. Hyperbolic Differ. Equ. **7** (2010), 605–649.
- [63] G. Wolansky, *Dynamics of a system of sticking particles of finite size on the line*, Nonlinearity **20** (2007), no. 9, 2175–2189.
- [64] E. H. Zarantonello, *Projections on convex sets in Hilbert space and spectral theory. I. Projections on convex sets*, Contributions to nonlinear functional analysis (Proc. Sympos., Math. Res. Center, Madison, WI, 1971), Academic Press, New York, 1971, pp. 237–341.
- [65] Ya. B. Zel’dovich, *Gravitational instability: An approximate theory for large density perturbations*, Astro. Astrophys. **5** (1970), 84–89.

FABIO CAVALLETTI, UNIVERSITÀ DEGLI STUDI DI PAVIA, DIPARTIMENTO DI MATEMATICA, VIA
FERRATA 1, 27100 PAVIA, ITALY

E-mail address: `fabio.cavalletti@unipv.it`

MARC SEDJRO, LEHRSTUHL FÜR MATHEMATIK (ANALYSIS), RWTH AACHEN UNIVERSITY, TEM-
PLERGRABEN 55, 52062 AACHEN, GERMANY

E-mail address: `sedjro@instmath.rwth-aachen.de`

MICHAEL WESTDICKENBERG, LEHRSTUHL FÜR MATHEMATIK (ANALYSIS), RWTH AACHEN UNI-
VERSITY, TEMPLERGRABEN 55, 52062 AACHEN, GERMANY

E-mail address: `mwest@instmath.rwth-aachen.de`